Characterising the host response to the emerging Ebola virus, Makona variant from West Africa

Andrew Bosworth

University of Liverpool

This thesis is submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Andrew James Bosworth

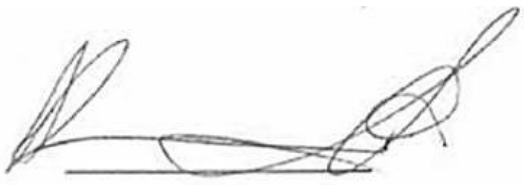
November 2014-November 2018

Author Declaration

I declare that this thesis was composed by myself alone, that the work contained herein is my own except where explicitly stated otherwise in the text and referenced appropriately, and that this work has not been submitted for any other degree or processional qualification except as specified. Any contributions from colleagues in collaboration, such as in the conduct of high containment operations, high security operations, or other activities restricted to specially trained individuals, are explicitly referenced in the text.

Work presented in Chapter 3 has been accepted for publication in Scientific Reports, under the title "*A comparison of host gene expression signatures associated with infection in vitro by the Makona and Ecran (Mayinga) variants of Ebola virus*", I carried out all laboratory activities, analyses, wrote the manuscript and in collaboration with my Supervisor *Julian A. Hiscox* formatted this work for publication.

Andrew James Bosworth



To Katherine

For your patience, support and encouragement

Acknowledgements

I thank Professor Roger Hewson, for the opportunities that have been afforded to me, advice, guidance and supervision. I thank Dr. Stuart D. Dowall for support and encouragement throughout this project, and for his efforts in performing intricate work at high containment. I thank Professor Julian A. Hiscox for advice, guidance and supervision. I thank Professor Miles W. Carroll for advice and insight throughout this project. I acknowledge the fantastic work of the High Containment Microbiology department, which provided support for the high containment activity throughout this work. I acknowledge the assistance of staff from the Centre for Genomics Research (CGR) at the University of Liverpool with analysis of transcriptomic data. I acknowledge the assistance and advice from Dr. Stuart Armstrong, for his help performing proteomic experiments and analysis and Dr. David Matthews, for his help performing complex bioinformatics analysis on viral genomes. The work described in this thesis was completed with financial support from the National Institute for Health Research, UK, as a scholarship and research support awarded by the Health Protection Research Unit for Emerging & Zoonotic Infections, a collaborative institution set up between the Institute of Infection & Global Health, University of Liverpool and the National Infections Service, Public Health England, UK. I dedicate this thesis to all those who fought the spread of Ebola Virus Disease in West Africa.

Abstract

Characterising the host response to the emerging Ebola virus, Makona variant from West Africa

By Andrew Bosworth

West Africa was in the grips of an Ebola Virus Disease outbreak, caused by the emerging Makona variant of Ebola virus. High resolution molecular methods including transcriptomics and proteomics were utilised to profile the host response to the emergent Makona variant from West Africa, and compare this response with that induced by infection with other *ebolaviruses,* in order to identify host factors potentially important in host pathology. A comparison between Makona and other well characterised variants of Ebola virus showed that induced differences in the host response were not significant (Chapter 3) and that the transcriptomic changes were very similar to previously characterised isolates. To evaluate the importance of interferon to the lifecycle of the Makona variant, *in vitro* comparisons with Reston virus were performed to highlight important changes in the antiviral state of multiple cell lines during infection, this showed an effective interferon response was not a major determinant of successful *ebolavirus* infection (Chapter 4). The pro-inflammatory response to the Makona variant and Reston virus were compared in a relevant inflammatory cell type (Chapter 5). Analysis indicated that a highly active NFκB response may be required for efficient virus replication, indicating a potent inflammatory response is essential for the virus lifecycle (Chapter 6). The Makona variant of Ebola virus was hypothesised to induce distinctive transcriptional and proteomic changes in infected cells. In this thesis, evidence is presented that infection with the Makona variant does not induce significantly different patterns of host response from that observed in other *ebolaviruses,* and presents the first longitudinal transcriptomic analysis of patient infected with the Ebola virus, Makona variant. Furthermore, this study has revealed the critical role of NFκB in the lifecycle of the *ebolaviruses*.

Associated Publications

The work performed in the completion of this thesis resulted in the publication of the following manuscripts created from data discussed in thesis chapters:

I. **Andrew Bosworth**, Stuart D. Dowall, Natasha Y. Rickett et al: *A comparison of host gene expression signatures associated with infection in vitro by the Makona and Ecran (Mayinga) variants of Ebola virus.* Scientific Reports 01/2017

Manuscripts related to contributions made to the European Mobile Laboratory in their diagnostic services and subsequent research in West Africa:

II. Romy Kerber, Ralf Krumkamp, Boubacar Diallo, … , **Andrew Bosworth** et al: *Analysis of Diagnostic Findings From the European Mobile Laboratory in Guéckédou, Guinea, March 2014 Through March 2015.* The Journal of Infectious Diseases 09/2016; 214(Suppl 3). DOI:10.1093/infdis/jiw269

III. Miles W Carroll, David A Matthews, Julian A Hiscox, … , **Andrew Bosworth** et al: *Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa.* Nature 06/2015; 27. DOI:10.1038/nature14594

Manuscripts related to contributions made to the testing of specimens in important public health research, allowing validation of molecular assays.

IV. Stuart D Dowall, **Andrew Bosworth**, Robert Watson et al: *Chloroquine inhibited Ebola virus replication in vitro but failed to protect against infection and disease in the in vivo guinea pig model.* Journal of General Virology 10/2015; 96(12). DOI:10.1099/jgv.0.000309

V. Stuart D. Dowall, **Andrew Bosworth**, Emma Rayner et al: *Post-exposure treatment of Ebola virus disease in Guinea pigs using EBOTAb, an ovine antibody-based therapeutic.* Scientific Reports 07/2016; 6. DOI:10.1038/srep30497

VI. Stuart David Dowall, Jo Callan, Antra Zeltina, Ibrahim Al-Abdulla, Thomas Strecker, Sarah K Fehling, Verona Krähling, **Andrew Bosworth** et al: *Development of a Cost-effective Ovine Polyclonal Antibody-Based Product, EBOTAb, to Treat Ebola Virus Infection.* The Journal of Infectious Diseases 12/2015; 213(7). DOI:10.1093/infdis/jiv565

VII. Stuart D. Dowall, Kevin Bewley, Robert J. Watson, Seshadri S. Vasan, Chandradhish Ghosh, Mohini M. Konai, Gro Gausdal, James B. Lorens, Jason Long, Wendy Barclay, Isabel Garcia-Dorival, Julian Hiscox, **Andrew Bosworth** et al: *Antiviral Screening of Multiple Compounds against Ebola Virus.* Viruses 10/2016; 8(12)., DOI:10.3390/v8110277.

Additional Complementary Work

The methodologies and scientific knowledge developed during the course of this project provided support to critical public health research, contributing to the following publications and manuscripts:

I. David N. Bukbuk, Stuart D. Dowall, Kuiama Lewandowski, **Andrew Bosworth** et al: *Serological and Virological Evidence of Crimean-Congo Haemorrhagic Fever Virus Circulation in the Human Population of Borno State, Northeastern Nigeria.* PLoS Neglected Tropical Diseases 12/2016; 10(12)., DOI:10.1371/journal.pntd.0005126

II. Stuart D. Dowall, Victoria A. Graham, Emma Rayner, Barry Atkinson, Graham Hall, Robert J. Watson, **Andrew Bosworth** et al: *A Susceptible Mouse Model for Zika Virus Infection.* PLoS Neglected Tropical Diseases 05/2016; 10(5). DOI:10.1371/journal.pntd.0004658

III. Fares Wasfi, Stuart Dowall, Tayssir Ghabbari, **Andrew Bosworth** et al: *Sero-epidemiological survey of Crimean-Congo hemorrhagic fever virus in Tunisia.* Parasite 03/2016; 23. DOI:10.1051/parasite/2016010

IV. F. Wasfi, K. Dachraoui, S. Cherni, **A. Bosworth** et al: *West Nile virus in Tunisia, 2014: First isolation from mosquitoes.* Acta tropica 03/2016; 159. DOI:10.1016/j.actatropica.2016.03.037

V. **Andrew Bosworth**, Tayssir Ghabbari, Stuart Dowall et al: *Serological evidence of exposure to Rift Valley Fever virus detected in Tunisia.* 10/2015; New Microbes New Infections. DOI:10.1016/j.nmni.2015.10.010

VI. Edmund N C Newman, Penelope Johnstone, Hannah Bridge, Deborah Wright, Lisa Jameson, **Andrew Bosworth** et al: *Seroconversion for Infectious Pathogens among UK Military Personnel Deployed to Afghanistan, 2008–2011.* Emerging infectious diseases 12/2014; 20(12). DOI:10.3201/eid2012.131830

Table of Contents

List of Figures and Tables

Glossary

ACDP ..... Advisory Committee for Dangerous Pathogens
AHF ..... African Haemorrhagic Fever
Ambic ..... Ammonium Bicarbonate
APC ..... Antigen Presenting Cell
BDBV ..... Bundibugyo virus
CAR. ..... Central African Republic
CDC. ..... Centres for Disease Control and Prevention
CFR ..... Case Fatality Rates/Ratio
CL ..... Containment Level
CPE ..... Cytopathic Effect
DC ..... Dendritic Cells
DMEM ..... Dulbecco's Modified Eagles Media
DMSO ..... Dimethyl Sulfoxide
DNA ..... Deoxyribonucleic Acid
DRC ..... Democratic Republic of the Congo
EBOV ..... Ebola virus
ECACC ..... European Collection of Authenticated Cell Cultures
EHF. ..... Ebola Haemorrhagic Fever
ELISA ..... Enzyme Linked Immunosorbent Assay
EVD ..... Ebola Virus Disease
EZI ..... Emerging and Zoonotic Infections
FASP ..... Filter Aided Sample Preparation
FBS ..... Foetal Bovine Serum
FC ..... Fold Change
FDR ..... False Discovery Rate
FPKM ..... Fragments per Kilobase per Million reads
GLM ..... General Linear Models
GP ..... Glycoprotein
HPLC ..... High Performance Liquid Chromatography
HPRU ..... Health Protection Research Unit
HSP ..... Heat Shock Protein
ICTV. ..... International Committee on Virus Taxonomy
IFN ..... Interferon
IFITM ..... Interferon Interacting Transmembrane Protein
IL ..... Interleukin
IPA ..... Ingenuity Pathway Analysis
IRF ..... Interferon Regulatory Factor
ISG ..... Interferon Stimulated Gene
L ..... Viral Polymerase
LC ..... Liquid Chromatographer/Chromatography
LDS ..... Lithium Dodecyl Sulfate
LIOV ..... Lloviu virus
logFC/log2FC ..... logarithmic Fold Change
LPS ..... Lipopolysaccharide
MARV ..... Marburg virus

MEM ..... Minimum Essential Media
MOI ..... Mode of Infection
MS ..... Mass Spectrometer/Spectrometry
MSF ..... Medicins Sans Frontiere
MTT ..... 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MVD ..... Marburg Virus Disease
m/z ..... Mass to Charge ratio
NBF ..... Neutral Buffered Formalin
NFκB ..... Nuclear Factor kappa-B
NHP ..... None Human Primate
NIHR ..... National Institute of Health Research
NP ..... Nucleoprotein
OMS ..... Organisation Mondiale de la Sante
ORF ..... Open Reading Frame
PBS ..... Phosphate Buffered Saline
PBMC ..... Peripheral Blood Monocytic Cells
PCA ..... Principle Component Analysis
PHE ..... Public Health England
PMA Phorbol-13-myristate
qRT-PCR ..... Quantitative Reverse Transcription-Polymerase Chain Reaction
RESTV ..... Reston virus
RdRp ..... RNA Dependent RNA Polymerase
RIPA ..... Radio-immuno Precipitation Assay
RNA ..... Ribonucleic Acid
RNP ..... Ribo-nuclear protein
SDS ..... Sodium Dodecyl Sulfate
sGP ..... Soluble/Secretory Glycoprotein
ssGP ..... Small Secreted Glycoprotein
SILAC ..... Stable Isotope Labelling with Amino acids in Cell culture
SHFV ..... Simian Haemorrhagic Fever virus
SUDV ..... Sudan virus
TAFV ..... Tai Forest virus
TEM. ..... Transmission Electron Microscopy
TF. ..... Tissue Factor
TLR ..... Toll-like Receptor
TNF ..... Tumour Necrosis Factor
UK ..... United Kingdom of Great Britain and Northern Ireland
US/USA ..... United States of America
UTR ..... Untranslated Region
VP ..... Virus Protein (40, 35, 30, 24)
WHO ..... World Health Organisation

1. Introduction

1.1 Project Synopsis

The research described in this thesis was conceived at a critical time. West Africa was in the grips of an Ebola Virus Disease (EVD) outbreak, caused by an emergent *ebolavirus* that was as yet uncharacterised. At the outset of this project there was little data on the immune pathology or molecular biology of infection caused by the Ebola virus (EBOV), Makona variant.

Transcriptomics and proteomics are the study of all transcripts or proteins produced by a cell, a tissue or an entire organism (1). The application of transcriptomics and proteomics to study the host response to *ebolavirus* infection will expand knowledge of the mechanisms underlying infection with emerging and re-emerging *ebolaviruses*. To create the body of research described in this thesis, high resolution molecular methods were utilised to profile the host response to the emergent Makona variant of EBOV from West Africa. Analysis of infection with the Makona variant of EBOV was compared with the host response to other variants and other species of *ebolavirus* in order to identify host factors potentially important in host pathology: Aiding in the design of novel approaches to treat patients suffering with EVD and highlighting new areas of research into viral biology.

A large body of *ebolavirus* research has developed in the 41 years since discovery (2). In this chapter this body of research on *ebolaviruses* was explored and explained, the significant gaps in knowledge have been identified and the purpose of this thesis outlined.

1.2 History

The earliest reports of filovirus infection of primates originated from Marburg an der Lahn, Germany in 1967. 31 human cases and 7 deaths were reported, in an outbreak caused by an unidentified organism that affected humans and primates in Marburg, Frankfurt and Belgrade (3, 4). The *Filoviridae* (order Mononegavirales) are a family of non-segmented, negative sense, enveloped RNA viruses (5), and have been confirmed as the aetiological agents of fatal haemorrhagic fevers in African non-human primates and humans, resulting in thousands of deaths over the last ~40 years (6). The virus family can be split into three genera, based on phylogenetic relationship; the *marburgviruses*, *cuevaviruses* and the *ebolaviruses* (5).

1.2.1 The *ebolaviruses*

The *ebolaviruses* are the best characterised and studied genus of the Filoviridae. The *ebolaviruses* have also caused large numbers of human cases (~30,000 as of July 2017 (7)). The *ebolavirus* genus contains five species of virus (5). A phylogenetic tree representing this genus is shown in figure 1.1.

The viruses can be listed by highest recorded case-fatality rate in humans to lowest: *Zaire ebolavirus* (Ebola virus; EBOV) (6), *Sudan ebolavirus* (Sudan virus; SUDV) (8), *Tai Forest ebolavirus* (Tai Forest virus; TAFV, formerly *Cote d'Ivoire Ebola virus*) (9), *Bundibugyo ebolavirus* (Bundibugyo virus; BDBV)(10) and *Reston ebolavirus* (Reston virus; RESTV) (11). Viruses of the *ebolavirus* genus were first discovered in 1976 (2, 8, 12, 13), when two outbreaks occurred simultaneously. One outbreak took place in Yambuku, Democratic Republic of the Congo (DRC; then Zaire) (2) and another taking place in Nzara, Sudan (now South Sudan) (8). Early investigations into these outbreaks determined that the disease in these outbreaks was caused by

Marburg virus (MARV) or a related virus (14). However several years of investigation determined that the Sudanese outbreak was caused by SUDV (8), and the Zaire outbreak was caused by EBOV (2).

1.2.1.1 Ebola virus (EBOV)

The first confirmed human case of Ebola Virus Disease (EVD) caused by EBOV was in 1976 (2, 12). Reports emerged from the Democratic Republic of the Congo (then Zaire) of a deadly illness spreading amongst the rural populations in the north. The epicentre of the outbreak was located in Yambuku, a small town in Equateur province, close to the Ebola river (15).

EBOV did not emerge again until 1995 when an Italian nurse died in Soeurs des Pauvres de Bergame, Kikwit, DRC (16, 17). This showed the 1976-1977 EVD outbreaks were not isolated events, and the virus was likely endemic in the region. This also coincided with 5 separate outbreaks of EVD in Gabon (1995) (18). Human illness caused by EVD is characterised by severe illness, where those who develop the severest form of disease suffer overt haemorrhagic manifestations, dysentery, haematemesis and delirium (6). Case fatality rates (CFR) have ranged from 50%-90% in most outbreaks (7, 15).

Advancement of sequencing and phylogenetic analysis has resulted in the defining of several EBOV variants, distinct from strains due to similarity of genotype as determined by ICTV (5, 19, 20). The first variant to be discovered was the Yambuku variant (2), Mayinga and Ecran isolates (21) from the DRC 1976 outbreak. There are also Kikwit, Gabon, Makokou, Etakangaye, Mendemba, Entsiami, Olloba, Mvoula, Bouee and Luebo (5), Makona (Guinea) (22) and the Lokolia variants (DRC) (23).

1.2.1.2 Sudan virus (SUDV)

The virus isolated from the Sudanese outbreak in 1976 was classified as a separate species of virus to EBOV, named SUDV. This virus causes a slightly lower CFR of between 25% and 75% (6). EVD caused by SUDV is indistinguishable to that caused by EBOV (13).

1.2.1.3 Bundibugyo virus (BDBV)

Bundibugyo (pronounced *boon-dee-boo-joh*) is a small western Ugandan town, ~200 miles from the capital of Kampala. Outbreaks of EVD caused by BDBV are unusual in that they occurred on the border between the DRC and Uganda (6, 10). CFR in outbreaks of BDBV is estimated at ~40%. EVD caused by BDBV is distinguishable from that caused by EBOV as <50% of cases show severe manifestations (10, 24).

1.2.1.4 Tai Forest virus (TAFV)

Only one human case of EVD caused by TAFV has been confirmed. The confirmed case was in a primatologist studying chimpanzees (*Pan troglodytes*) in the Tai Forest National Park in Ivory Coast (*Cote d'Ivoire*) (6, 9). TAFV mostly remains uncharacterised due to few spill-over events being recorded. TAFV is believed to be biologically similar to other *ebolaviruses*, though further study may contradict this.

1.2.1.5 Reston virus (RESTV)

Human infections with RESTV have never been confirmed by either molecular analysis or viral isolation (24, 25). No symptoms are elicited in humans with serological evidence of exposure (26). The first outbreak of RESTV occurred in a primate facility in Reston, Pennsylvania, US. The primates were co-infected with simian haemorrhagic fever virus (SHFV) and RESTV; meaning the cause of severe symptoms in these primates is difficult to

classify (24, 27, 28). Later outbreaks in primate populations (mainly in captivity) in Russia, the US, Italy, Germany and the Philippines have reported lower levels of mortality and disease severity compared with initial descriptions (27). A large outbreak was also discovered in pigs in China, alongside further evidence of human exposure (29). The zoonotic transmission risk of RESTV means it is critical to better understand the danger this virus poses to human health (25).

1.2.2 Outbreaks

There have been more than 39 incidents and outbreaks involving *ebolavirus* infections since first discovery (7). Human CFR in *ebolavirus* outbreaks reach from 0% (TAFV, RESTV) to 90% (EBOV) (24). A summary of all outbreaks and incidents is shown in Table 1.1. Confirmation of outbreaks requires access to adequate testing facilities or the ability to transport samples to them. Additionally, sporadic cases have emerged outside of outbreaks, including laboratory infections. One such case occurred around the time of the outbreaks in Zaire and Sudan in 1976-1977 in the UK (12). In this incident a laboratory worker accidently inoculated himself with SUDV, and 6 days later suffered severe symptoms. The patient recovered with high intensity life support (12). Since this event there has been a further 7 reports of laboratory associated infections with an *ebolavirus* (30).

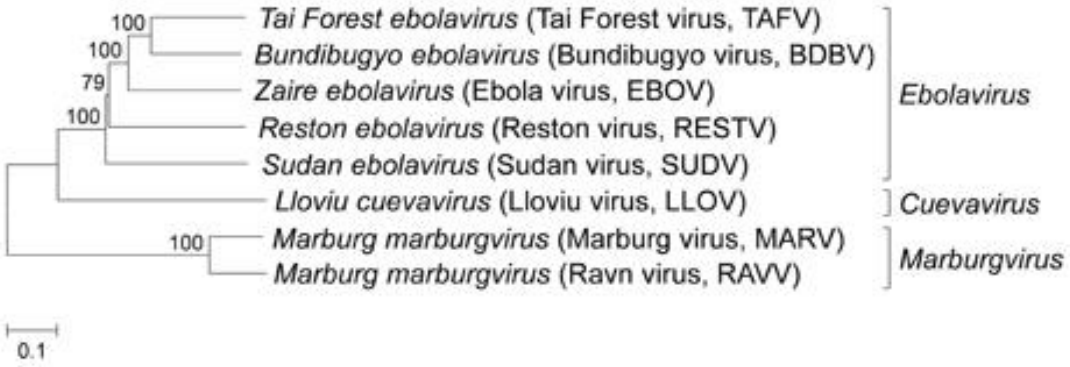


Figure 1.1: Representative neighbour joining phylogenetic tree of the Filoviridae, including ebolaviruses, cuevaviruses and marburgviruses. This phylogenetic tree was reproduced from Nakayama E and Saijo M. 2013. Animal models for Ebola and Marburg virus infections. It highlights the close relationship of Ebolaviruses and the clustering separately from Marburg virus and Lloviu virus. The scale of the analysis is 0.1 substitution per thousand nucleotides. GenBank accession numbers of the viral sequences used I nte construction of this tree are FJ217162, FJ217161, AF086833, AB050936, AY729654, JF828358, DQ217792, and EF446131.

Year	Country	Human cases	Deaths	
Ebola virus (EBOV)				
2017	Democratic Republic of the Congo	8	4 (50%)	
2014	Democratic Republic of the Congo	66	49 (74%)	
2014	Guinea, Liberia, Sierra Leone	28652	11325 (40%)	
2008	Democratic Republic of the Congo	32	15 (47%)	
2007	Democratic Republic of the Congo	264	187 (71%)	
2004	Russia	1	1 (100%)	Laboratory
2003	Republic of the Congo	35	29 (83%)	
2003	Republic of the Congo	143	128 (89%)	
2002	Republic of the Congo	57	43 (75%)	
2002	Gabon	65	53 (82%)	
1996	Russia	1	1 (100%)	Imported
1996	South Africa	2	1 (50%)	
1997	Gabon	60	45 (74%)	
1996	Gabon	37	21 (57%)	
1995	Democratic Republic of the Congo	315	250 (81%)	
1994	Gabon	52	31 (60%)	
1977	Zaire	1	1 (100%)	Isolated Case
1976	Zaire	318	280 (88%)	
Sudan virus (SUDV)				
2012	Uganda	6	3 (50%)	
2012	Uganda	11	4 (36.4%)	
2011	Uganda	1	1 (100%)	Imported
2004	Sudan (South Sudan)	17	7 (41%)	
2001	Uganda	425	224 (53%)	
1979	Sudan (South Sudan)	34	22 (65%)	
1976	United Kingdom	1	0	Laboratory
1976	Sudan (South Sudan)	284	151 (53%)	
Bundibugyo virus (BDBV)				
2012	Democratic Republic of the Congo	36	13 (36.1%)	
2007	Uganda	149	37 (25%)	
Tai Forest virus (TAFV)				
1994	Côte d'Ivoire (Ivory Coast)	1	0	Field Study
Reston virus (RESTV)				
2008	Philippines	6	0	
1996	Philippines and USA	0	0	Laboratory
1992	Italy	0	0	Laboratory
1989	Philippines	3	0	
1990	USA	4	0	Laboratory

Table 1.1: Summary of outbreaks and incidents where EVD was confirmed. Location of incidents and outbreaks is shown alongside the year of each event. The number of human cases/deaths is shown with percentage CFR. This data was sourced from the CDC accessed: 05/06/2017 (30). Cases with substantiated relatedness to laboratory or isolated incidents such as disease importation or field studies are indicated on the right of the table.

1.2.3 Zoonoses

A zoonosis is a pathogen which is transmissible from an animal host to humans. Zoonotic viruses may cause disease in humans but asymptomatically infect the reservoir animal host. The zoonotic potential of filoviruses has been known from first discovery (2, 3, 8, 31). The interest in bats (32) as a potential reservoir (32) stems from the epidemiological link of a MVD case in a research student in Rhodesia (Zimbabwe) with possible exposure while visiting the Sinoia Caves (33, 34). Additionally all cases associated with outbreaks of MVD in the DRC between 1998 and 2000 were associated with illegal gold mining (35). Following the identification of human MVD cases associated with the Kitaka Caves in Uganda, local bat species were investigated. The confirmation of MARV in *Rousettus aegypticus* bats (36), meant future research on finding the reservoir of *ebolaviruses* was focussed on bat species. Genetic material from *ebolavirus*es (RESTV) was recovered from bats in South-East Asia (37, 38) and from bats in Gabon (32). Serological reactivity was also found in bats and other forest-dwelling terrestrial mammals (31, 32, 39).

1.2.3.1 Bats

Experimental inoculation of *Tadarida condylurus* (Angolan free-tailed), *Tadaria pumila* (little free-tailed) bats have resulted in productive infection (31). Reactive antibodies and EBOV viral RNA was detected in the spleens and livers of *Hypsignathus monstrosus, Epomops franqueti* and *Myonycteris torquata* (figure 1.1) (32, 40) and high prevalence of EBOV reactive antibodies was found in these bat species (41). All three species are part of the human diet in Gabon, the Congo and the DRC (32, 42). Distribution of all three bat species overlaps closely with the distribution of cases in Africa (40).

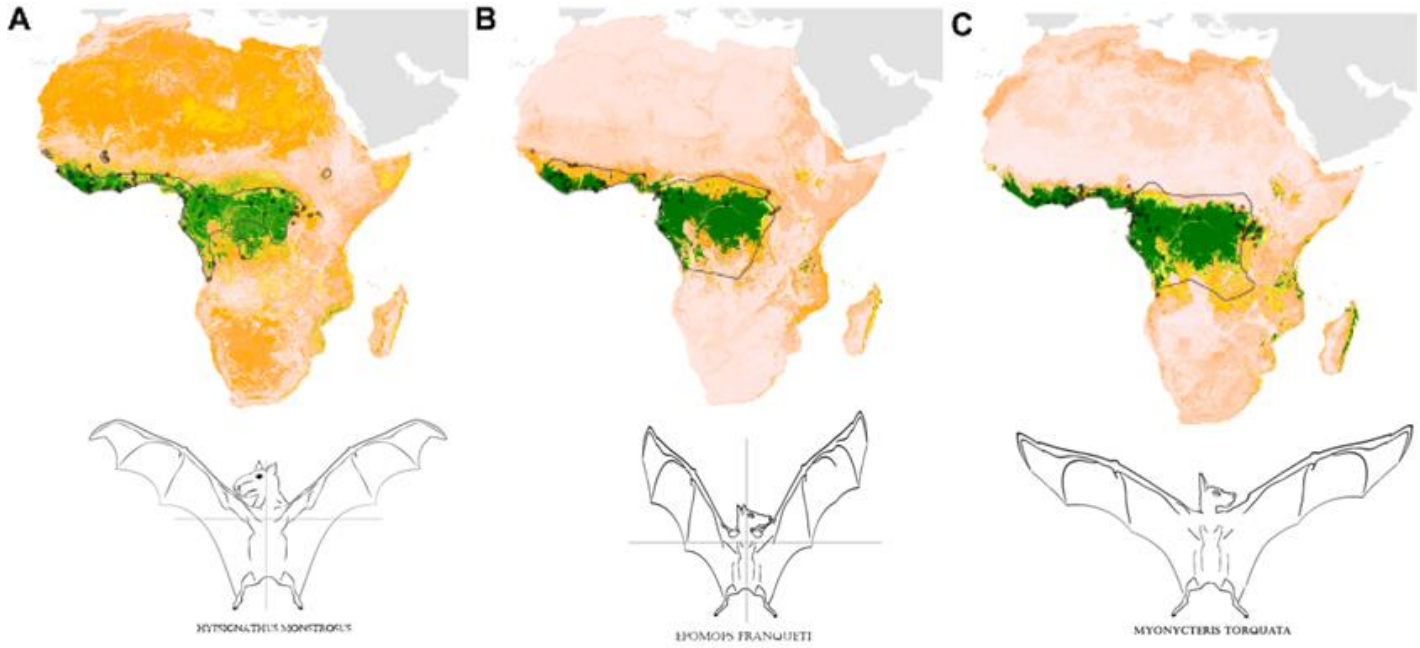


Figure 1.2: Distribution of bat species implicated as ebolavirus reservoirs in Africa. (A) Distribution of Hypsignathus monstrosus, a fruit eating species of megachiroptera indigenous to West and Central Africa. (B) Distribution of Epomops franqueti. (C) Distribution of Myonycteris torquata. All three species are frugivorous (fruit eating) and roost in high canopies in forested areas, usually away from human settlement. The distribution maps show similar concentrations of these species in central Africa (CAR, DRC, Congo, Gabon) and in West Africa (Liberia, Sierra Leone, Guinea, Cote d'ivoire, Nigeria). A scale bar is provided to indicate the relative abundance of bats found in each area, white indicating no bats found, and green colour indicating a high abundance of bats in these areas. Bat images are hand-drawn representations based on images collected online and photographs. The map images and distribution data were adapted from Pigott et al (40).

Figure 1.3: Photograph of the terrain surrounding Guékédou, Guinea near to a tributary of the Moa River. This area is known as Guinea Forestiere (Forested Guinea); large scale deforestation follows a kimberlite vein through the nearby hills, where diamond mining industry expanded. Thus the natural habitat of forest dwelling mammals was removed, and trees are much sparser close to the settlement. This photograph was taken less than 3 miles from the treatment centre at the epicentre of the outbreak of EBOV, Makona in July 2014.

All three species roost high in canopies, rarely coming into contact with humans except when predated or human activity disrupts habitats (32, 40). Large scale deforestation has taken place in sub-Saharan Africa, clearing the way for lucrative mining industry (gold, copper, tin and diamonds), farming, and expanding urbanisation (43). Disruption of bat habitats causes them to migrate to other areas, bringing them closer to humans and removing the barriers which ordinarily impede transmission. Some of the geographical changes introduced by humans are visible in Guinea, where EBOV re-emerged in 2013/14. Between 1990 and 2005 Guinea lost ~9% of its forest cover (44). figure 1.2 is a photograph of Guékédou, previously a forested region.

1.2.4 Transmission

The principle transmission routes of *ebolaviruses* are by direct contact with contaminated body fluids and fomites (45). Aerosol transmission of *ebolaviruses* is possible under laboratory conditions (46, 47). An unsettling characteristic of EVD is post-mortem transmission. Bodies of those whom have died from EVD continue to secrete virus after death (48), this makes funerals involving traditional practices particularly dangerous. WHO and ACDP clinical guidelines prohibit post-mortem of EVD patients due to the inherent risks. This has complicated studies of tissue impact in patients, with the only study having been performed on biopsied samples collected during the first outbreak in 1977 [68]. Indeed human behavioural and social aspects may be more responsible for the transmissibility of EBOV in outbreaks than virus biology. Many human cultures affected by EVD engage in cultural practices which involve close contact with the deceased for religious reasons. During EVD outbreaks this leads to transmission within families and should be prevented by safe dignified burial (figure 1.3) (49).

Figure 1.4: Photograph of a dignified burial team from the International Committee of the Red Cross working in Sierra Leone. These teams were found to be essential for effective control of EVD transmission during the outbreak. Sourced from the CDC Public Health Image Library accessed: 06/07/17

Humans are not the only species at risk of filovirus infection. From 1994 to 2003 an estimated 44 epizootic outbreaks occurred in primate populations of the Congo and Gabon (50). Carcasses of 50 lowland Gorillas, 15 chimpanzees and 6 other unidentified species of primate were found in these epizootics (50, 51). Further review of areas affected by these outbreaks has estimated that ~5000 primates have died from filovirus infection since these viruses were discovered in 1968 (52). Non-human primates (NHP) therefore represent a sentinel of filovirus circulation, and an early warning that a potential outbreak in humans may occur in that region if not contained.

Hunting or poaching by humans also results in transmission by bringing humans into close proximity with infected primates, and are believed to have been responsible for outbreaks of EVD in Cote d'Ivoire and the Congo (9, 53, 54). The term "Spillover" is used when discussing the transmission of zoonoses between one species and another. This relationship is summarised in figure 1.4, which displays the interplay between suspected reservoir transmission to sentinel species and potentially humans from species which they previously would not have come into close contact. Interestingly, susceptible species which exhibit caring behaviour (primates) appear most vulnerable to transmission, where EBOV appeared to thrive in the ecological niche created by the compulsion to care for the sick.

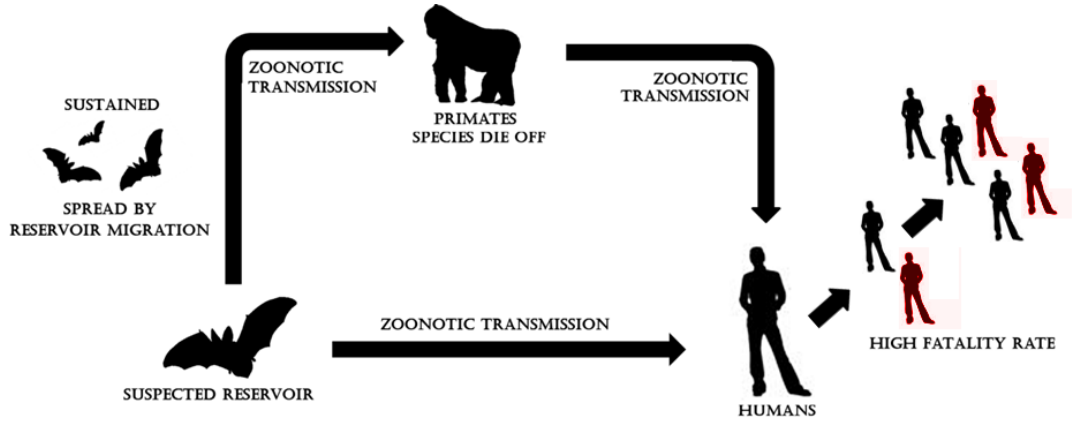


Figure 1.5: Representation of Filoviridae zoonotic transmission from suspected reservoir species (megachiroptera) and terrestrial mammal species. Principal transmission routes are shown through predation of the reservoir species by non-human primates, and either predation or other interaction by Homo sapiens. Transmission to primates, including humans, results in high casualty rate. Outbreaks in wild populations of NHP are typically detected due to mass die off in troops. Evidence suggested that ebolaviruses are maintained in the suspected reservoir populations and are distributed geographically through mass migration patterns. Other species not shown here are small terrestrial animals such as forest deer, believed to also play a role in transmission. Mass die off in non-primate hosts has not been detected. This is likely because primate species are rare in that they display care behaviour and live in close groups, leading to rapid spread of ebolaviruses in family units.

1.2.5 Re-emergence in West Africa

The re-emergence of EVD in West Africa was initially suspected to be malaria (22). The eastern forested region of Guinea is often plagued by malaria, cholera and other infectious diseases of concern. *Medicins sans frontiere* (MSF) had an existing malaria treatment station in Guékédou, Guinea. The doctors in this remote outpost initially sought clinical advice from their chief epidemiologist when a pattern of familial death was observed, not consistent with the diagnosis of malaria or cholera (7, 55). A treatment and isolation centre was established; the design of this centre (figure 1.5) was identical to that used in outbreaks of EVD and MVD in 2009 and 2012.

The first cases of EVD in the West African outbreak were reported in March 2014 and the first responders arrived 72h later (22, 55). Phylogenetic analysis suggested the Makona variant of EBOV entered the human population some time in December 2013 (56), this analysis is described in more detail in Chapter 3. The outbreak in West Africa began when a young boy was exposed to EBOV when interacting with a dead bat (55). By July 2014 EVD cases had been reported in frontier settlements in Sierra Leone (Kailahun), Liberia (Lofa) and Guinea (Guékédou, Macenta and Kissidougou), and in the capital of Guinea, Conakry.

Figure 1.6: Photograph of Ebola Treatment Unit, Guékédou, Guinea, taken in July 2014. Displayed is the main entrance to the MSF treatment unit with outer barrier. To the rear is the tent where the hot zone was located, housing confirmed EVD patients.

A map of case distribution was provided by the WHO throughout the outbreak. Figure 1.6 displays case distribution in the principally affected countries from March 2014 until March 2015. In total, approximately 28652 cases were recorded in what is now known to be the largest outbreak of EVD ever recorded (7). >11000 people lost their lives to EVD (7, 57).

1.2.5.1 Scale of the outbreak

Guékédou is in one of the remoter parts of Guinea, near the borders with both Sierra Leone and Liberia. All three countries are former colonies of European powers, the border between them porous and doing little to control the flow of people, who have cultural, family and religious homology.

This presented a major challenge for outbreak control, as it meant the flow of people was difficult to regulate, and quickly cases emerged across the border in Kailahun, Sierra Leone. Unlike all other previous outbreaks the affected area was not isolated; travel and road conditions were reasonable. Travel between major urban population centres and the outbreak zone were unimpeded. Disturbingly, cases began to emerge in the capitals of Conakry, Monrovia and Freetown (57). Whether or not differences in the biology of the West African variant of EBOV could contribute to the outbreak scale was the subject of investigation during the project described in this thesis.

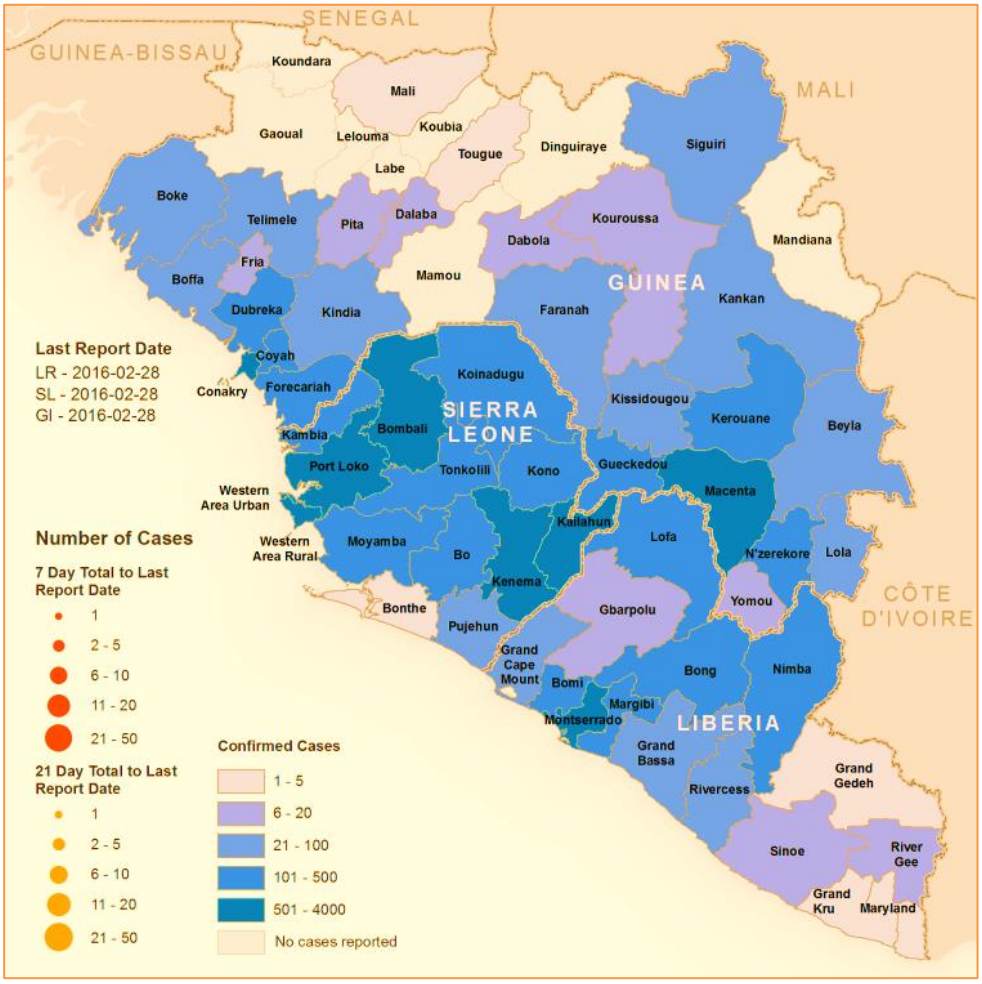


Figure 1.7: Map of Guinea, Liberia and Sierra Leone, the three countries affected by the EVD outbreak in 2013-2016. Provinces are coloured by number of cases reported between March 2014 and March 2015, the time period where samples were collected for analysis. This Ebola Outbreak Map [3/14-3/15] was sourced from World Health Organisation accessed: 05/06/2017 (7)

1.3 Virology

1.3.1 Virion structure

An *ebolavirus* virion has a filamentous morphology, forming particles uniformly ~60nm in diameter but which vary in length between 800nm and 1µm in length (14, 58, 59). The main structural proteins of EBOV are the membrane protein (VP40) and the glycoprotein (GP) (59). Sequestered within the EBOV virions are other proteins with roles in viral replication and viral transcription. VP35 and VP30 regulate these processes (60, 61); VP24 has many diverse roles in viral packaging (62), and as an antagonist of host restriction factors (63). The viral polymerase (L) is principally designed to replicate viral genome and transcribe genes to form new viral capsids (61). In order to package viral RNA within the virion the viral nucleoprotein (NP) binds to viral RNA and encapsulates it like a sheath, forming the ribonucleoprotein complex (RNP). This process packages the viral RNA and facilitates replication processes (64). The NP also protects viral RNA when initially infecting a host cell. A diagram of a viral particle is given in figure 1.7. Ultra-structural analysis of the recently isolated (C05 isolate) Makona variant from Guinea, showed that the morphology of this new virus is the same as previous EBOV isolates. The ultra-structure of the Makona variant is shown in figure 1.8.

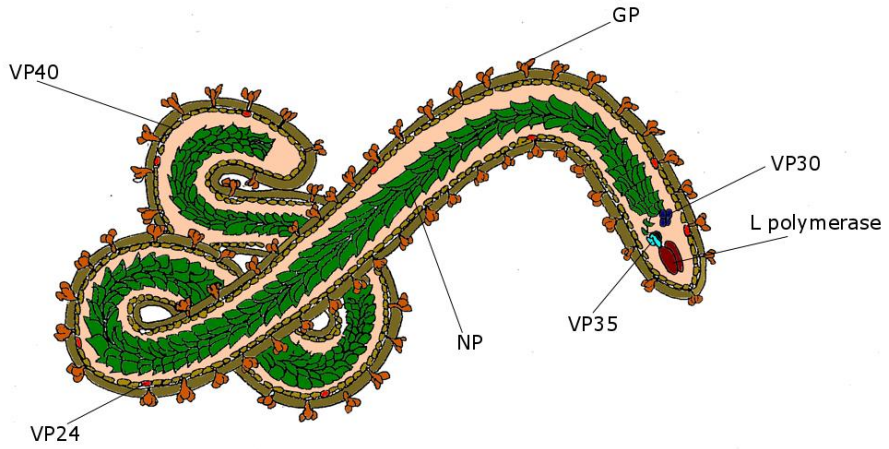


Figure 1.8: Hand-drawn depiction of a filovirus. The filamentous morphology gives name to the genus. The outer membrane is composed of interlinking polymeric structures called VP40. Embedded in this membrane is the viral fusion glycoprotein (GP). VP24 is associated partially within the membrane structure, and partially with the nucleoprotein, and has multiple roles in viral packaging and an antagonist of host restriction. Sequestered within the capsid the viral RNA is complexed with a polymeric protein structure known as the nucleoprotein (NP) to form the ribonucleoprotein (RNP) complex. The viral polymerase (L), transcriptional activator (VP35) and replication activator (VP30) are also shown, ready to begin viral replication upon entry into the host cell cytoplasm.

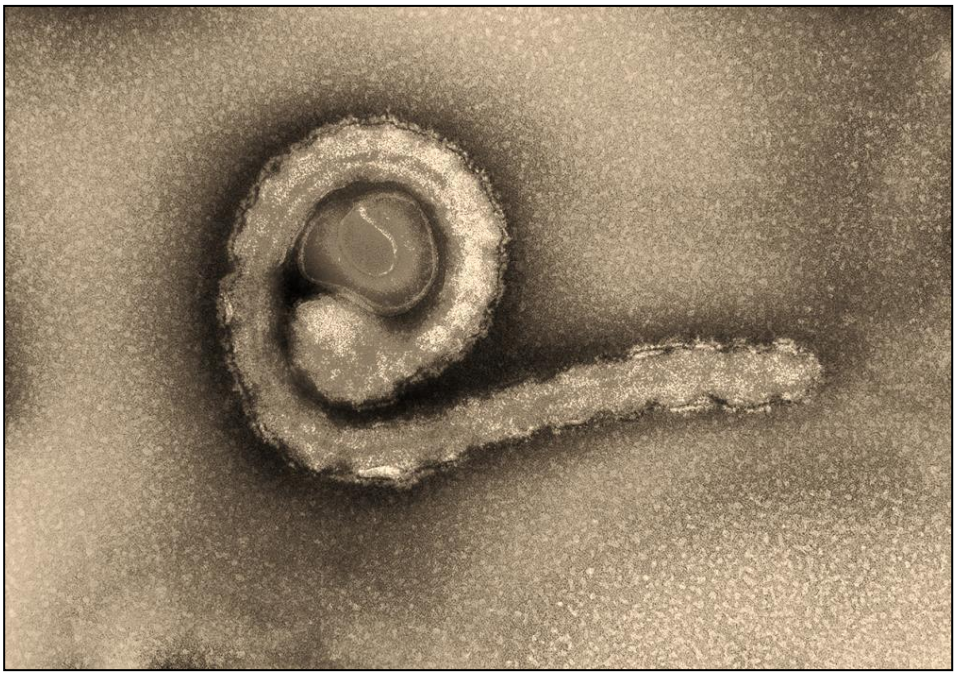


Figure 1.9: Transmission Electron Microscope (TEM) image of Ebola virus, Makona variant present in supernatants collected from infected in vitro cultured Vero cells for 72h. Shown in the field is the full length of an Ebola virus particle, with areas where the plane intersects a perpendicular contusion in the virion filamentous structure. Images were courtesy of Dr. Matthew Hannah, PHE and cultures were performed by Dr. Robin Gopal et al, PHE, permissions are shown in Appendix I.

1.3.2 Molecular virology

The *ebolavirus* 18.9Kb negative sense RNA genome encodes 7 genes flanked by inter-genic regions (65). From the 3' end transcription is sequential, beginning with the nucleoprotein (NP), listed sequentially are the transcription initiator (VP30), and replication co-factor (VP35), glycoprotein (viral fusion protein) (GP), viral matrix protein (VP40), primary host restriction factor antagonist (VP24) and viral polymerase (L) (62). 6 of these genes result in single transcriptional products while the GP gene produces 4 which have so far been identified (66, 67). These additional GP transcriptional products are regulated by an editing site located mid-way in the GP ORF (68, 69). The primary product of the GP gene is known as secretory GP (sGP), which is transported to the cell surface and released into the extracellular space (67). This also results in production of a smaller product known as ssGP (small secretory GP) (67), and a product known as Delta-Peptide (Δ peptide) (70, 71). A schematic diagram of the genome is shown in figure 1.9. Transcription and replication begins at the 3' initiation site encoded in the 3' UTR (untranslated region) (62). Models indicate that the polymerase sequentially transcribes the genes in the 3' to 5' direction (61, 72), transcription of individual genes is initiated at the start site located in the inter-genic region 3' of the encoding ORF (73), and regulated by the stop site in 5' of the ORF (figure 1.10).

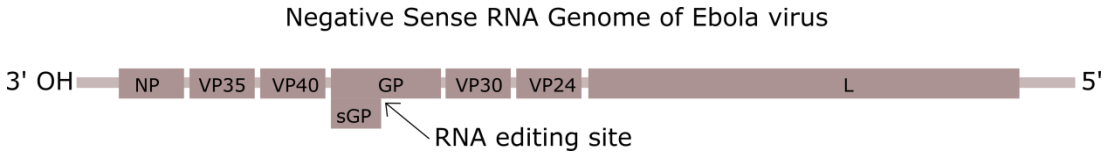


Figure 1.10: Schematic diagram of ebolavirus genomic architecture. All ebolaviruses are unsegmented negative sense RNA viruses. The genome is arranged 3' to 5' in the direction of replication and transcription as follows: NP, VP35, VP40, GP, VP30, VP24 and L. The RNA editing site controlling the transcription of soluble GP (sGP) and other GP gene products is annotated with an arrow. The viral gene involved in viral RNA encapsidation (NP) is located at the terminal 3' region, meaning this gene is most transcribed. VP35 is an accessory protein modulating NP binding, and is thus transcribed next in sequence. VP40 and GP encode structural proteins, and are next transcribed. VP30, VP24 and L are later in the sequence, and are non-structural proteins important in viral replication and transcription. Many RNA virus genomes follow this organisation. During replication full length viral genome is transcribed to form an anti-genome, which acts as a template for the production of viral progeny. Transcription of genes is in the 3' to 5' direction and was hypothesised to follow a step-wise gradient, with the greatest transcription at the 3' of the genome and the lowest levels of transcription at the 5' end. Image adapted from and sourced from the Ebola virus entry on viralzone.expasy.org accessed 01/07/2017 and again on 07/07/2017

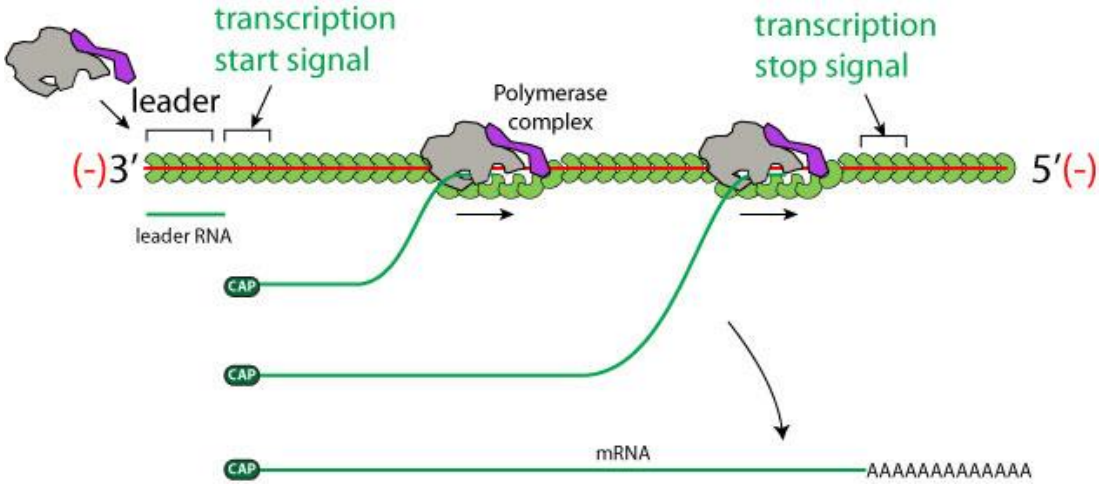


Figure 1.11: Schematic diagram of hypothesised replication and transcriptional strategy utilised by EBOV and other negative sense RNA viruses. In this model viral genes are sequentially transcribed by the viral polymerase in a step-wise fashion, forming a gradient of transcription across the genome from 3' to 5'. This gradient of transcription is thought to be shared in all Mononegavirales. The transcription of each gene is regulated by initiation (start) and stop sites 3' and 5' of the viral gene in extra genic regions. This means that the genes closer to the 3' start site are transcribed more regularly and transcript from these genes accumulate at greater abundance. Image adapted from and sourced from the Ebola virus entry on viralzone.expasy.org accessed 01/07/2017 and again on 07/07/2017

1.3.2.1 Replication

Replication of *ebolavirus*es occurs inside inclusion bodies (74), inclusion bodies are formed from an aggregation of *ebolavirus* nucleoprotein (NP) and host proteins. The main activity of the viral L polymerase occurs inside inclusion bodies, and synthesised ribonucleoprotein (RNP) complexes form, they migrate to VP40 (matrix protein) assemblies at the cytoplasmic surface.

1.3.2.2 Nucleoprotein (NP)

The viral nucleoprotein binds to viral RNA, encapsulating and forming the ribonucleoprotein complex (RNP) (75). This encapsulation shields RNA from degradation by cellular nucleases and facilitates transcription and replication (76). Absence of nucleoprotein makes replication impossible (62, 73). Crystallographic studies show that NP oligomerises into helical filaments around viral RNA (77), the formation of NP oligomers is partly regulated by phosphorylated VP35 (77, 78). Removal of VP35 causes nucleocapsid self-assembly and RNA binding (77, 78).

1.3.2.3 Viral Protein 30 (VP30)

VP30 is a viral RNA binding phosphoprotein (60), also called the minor nucleoprotein, which has a key role in transcription initiation (60, 79). Deletion of VP30 from the viral genome prevents transcription (62, 80), meaning this protein is essential for virus lifecycle.

VP30 forms homohexameric formations which bind to RNA and facilitate L polymerase activity on genes to produce transcripts for viral proteins (80).

1.3.2.4 Viral Protein 35 (VP35)

VP35 is a phosphoprotein which forms homotetrameric formations (80), binding to NP to regulate self-assembly, and thus regulating viral replication (78). VP35 also regulates viral replication by acting as a co-factor for the viral

L polymerase to enhance activity (73). VP35 also acts on cellular proteins to prevent the cell taking on an antiviral state by interfering with RNA sensing factors such as RIG-I and interferon regulatory proteins such as IRF-3 (81-83).

1.3.2.5 Glycoprotein (GP)

The GP gene encodes a soluble non-structural glycoprotein called sGP but is capable of producing structural GP1 and GP2 after editing in the gene transcript. GP2 is the fusion protein for *ebolaviruses* enabling host cell entry (59, 84, 85). GP1 forms a trimer anchored together by GP2, which mediates membrane fusion (86). A number of cellular receptors for GP have been identified including DC-SIGN (87, 88), LSECtin (89), hMGL (87), β-integrins (90) and the Tyro3 receptors (91). However deletion of any one of these receptors does not inhibit EBOV host cell entry and the presence of only one of these receptors is not sufficient to permit infection.

Therefore a critical receptor for EBOV host cell entry has yet to be identified. Entry into the cell is via endocytosis (91), resulting in the virion being trapped inside an endosome within the cell cytoplasm. To permit escape into the cytoplasm and cell infection GP must be catalysed by Cathepsin B/L, trimming the GP1 protein to a fraction of its size, removing the large mucin domain and revealing alternative binding sites for additional receptors within the endosome (92). Furin then catalyses a conformational change enabling GP to bind NPC1, the only identified critical entry factor for *ebolavirus*es (85). NPC1 is located on the inner membrane of endosomes and is involved in the trafficking of cholesterol (93). However the role of cholesterol biosynthesis and trafficking in EBOV infection has yet to be fully explored.

1.3.2.6 sGP

The GP gene primary product is sGP, a non-structural secretory glycoprotein. sGP is translated from un-edited transcripts from the GP gene (67, 94). GP1 and GP2 transcripts are produced from the GP gene through an editing event. A site containing 7 uracils (7U) are post-transcriptionally edited, introducing an adenosine insertion (A) in the transcript of this region, resulting in a frame-shift in the ORF. After editing the section of GP transcript encoding the C-terminal transmembrane region is translated, producing the transmembrane GP1 and GP2 (67). In unedited sGP transcription, the C-terminal region is also produced but is cleaved in the golgi (95), and is secreted from the cell (67, 94). The cleaved C-terminal region is also called delta (Δ) peptide, delta peptide is able to compete with EBOV particles for host cell surface receptors (70), modulating EBOV cell fusion, and may act as a viroporin, disrupting cellular membranes and facilitating viral egress (70). The N-terminal region of sGP, ssGP and GP1 are identical in structure and sequence for 295 amino acids; only the C-terminal region is different (95). Therefore the principle role of sGP may be as a glycosylated decoy, competing with full length GP for binding by reactive antibodies (86, 94). In this way reducing available neutralising antibody and assisting EBOV intra-host spread. sGP is also able to bind to available GP1 to form a heterotrimeric form of transmembrane bound complete GP indistinguishable from that composed of a homotrimer of GP1 bound to GP2 (67, 95). sGP is also hypothesised to activate endothelial cells, weakening cell junctions and exacerbating vascular leakage observed in EVD (71, 96). Occasionally 2 adenosines are inserted at the editing site, resulting in a truncated protein called ssGP, a smaller version of the soluble glycoprotein (67). A depiction of GP editing is provided in figure 1.11.

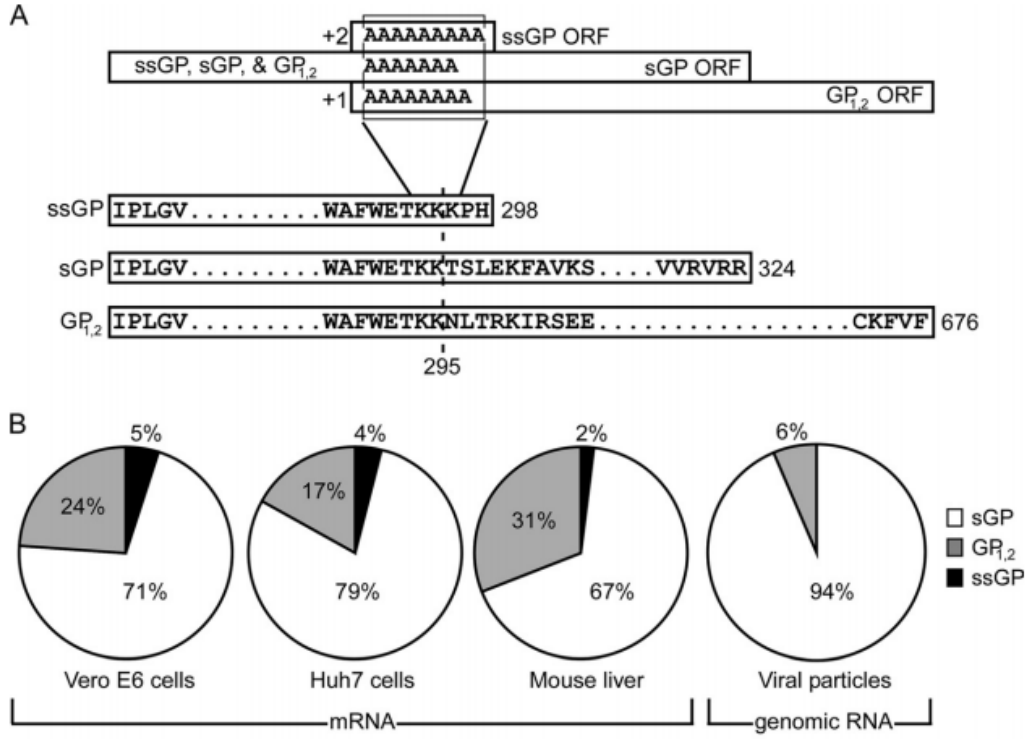


Figure 1.12: The editing of the ebolavirus GP gene. (A) This schematic diagram depicts the editing of the GP gene to create transcripts encoding ssGP by the insertion of 2 additional adenosines, or the creation of transcripts encoding GP1,2 by addition of 1 adenosine. This frame-shifts the encoding transcript to the ORF of sGP (7A/U), ssGP (9A/U) or GP1,2 (8A/U). The N-terminal region of 195 amino acids in each of these gene products is identical; a highly glycosylated region acting as fusion protein for viral cell surface receptor binding. (B) The proportion of sGP, GP1,2 and ssGP produced in Vero E6 cells, Huh7 (human liver derived) cells and mouse liver. Also shown is the amount of genomic RNA encoding sGP and GP1/2 in viral particles. This shows that the largest proportion of GP transcripts encode sGP in vitro and in vivo. An analysis of the proportion of sGP encoded by viral genomic RNA is evidence that sGP is the primary gene product of the ebolavirus GP gene. Schematic diagram adapted from Mehedi et al (2011) (67).

1.3.2.7 Viral Matrix Protein 40 (VP40)

The VP40 protein is an essential component in virion assembly (59, 97). The protein oligomerises to form stable structures at the cell plasma membrane (97). These are most concentrated in areas known as lipid rafts. Lipid rafts are areas of the plasma membrane with a particularly dense concentration of cholesterol and sphingolipids. Both cholesterol and sphingolipids form major components of the EBOV envelope after budding (98). VP40 initiates virion self-assembly at these lipid rafts, complexing with RNP to form complete virions and budding through the membrane, carrying envelope lipids and proteins associated with the infected host cells (59, 97).

VP40 also binds to the transmembrane C-terminal region of GP, anchoring it in place while its exposed mucin domain penetrates through the surrounding envelope (59). The role of lipid rafts in EBOV lifecycle is essential (98), and the apparent close involvement of virus lifecycle with sterol trafficking make this an attractive target in EBOV infection. VP40 has an additional role which is poorly understood. Large octomeric structures of VP40 have been observed in mammalian cells infected with EBOV, which appear to bind RNA and have a critical role in virion assembly (99), however the exact mechanisms of this are not understood.

1.3.2.8 Viral Protein 24 (VP24)

VP24 plays a prominent role in virus lifecycle as an antagonist of the host immune response and as a co-factor in virion assembly. VP24 associates with the RNP complex and has a suggested role in the packaging of NP inside the matrix and envelope structures (100), deletion of VP24 reduces virion packaging efficiency but does not interrupt the virus lifecycle (63, 100). Studies with mini-replicons have determined that VP24 had a negative effect on replicon activity (62). Mini-replicons are small in size and do not require VP24 for efficient packaging, while full length viral genomes with their associated NP structure does (100).

The arguably more important role for VP24 in virus lifecycle is its interferon antagonism (63, 101). VP24 binds to karyopherin-α, preventing the translocation of STAT1/STAT2 into the nuclear compartment to permit transcription of interferon response genes (102). This activity is conserved in all identified *ebolavirus*es including RESTV (63). VP24 is suggested to be a critical in host adaptation in guinea pigs (103), and sequencing of serially passaged virus from guinea pig infections found that genome mutations in this region accumulated when adapting to guinea pigs, correlating with increasing pathogenicity (104). Thus VP24 is a critical factor in determining viral pathogenicity (63, 101, 105).

1.3.2.9 Viral Polymerase (L)

The viral L protein is a large 258kDa RNA dependent RNA polymerase (RdRp) (73). The function of this protein is to associate with viral RNA, to transcribe viral mRNA (messenger RNA) and facilitate replication. The L polymerase binds to the 3' initiation domain by targeting a specific 3 dimensional RNA structure located in the 3' UTR (62, 73).

The processivity of L is tailored to either produce mRNA from viral genes by association with the viral phosphoprotein VP30, or producing full length viral RNA anti-genomes, regulated by concentration of NP (78). As NP is the first viral protein produced (and hypothetically the most abundant) and this molecular switch ensures NP is abundant enough to bind and protect the generated EBOV genomes (76, 78).

In this replicase configuration the L ignores all inter-genic start and stop signals and reads through all transcriptional sites and ends replication with an uncapped 5' untranslated region (106, 107). The viral L is highly conserved in all *ebolavirus*es due to its essential function in viral lifecycle, and is closely related in sequence and function to the RdRp of Mononegavirales (73).

1.3.3 Functions of viral proteins in *ebolavirus* lifecycle

All of these viral proteins perform defined tasks within the viral lifecycle from entry, replication and egress. The roles of these proteins in *ebolavirus* lifecycle is summarised in figure 1.12.

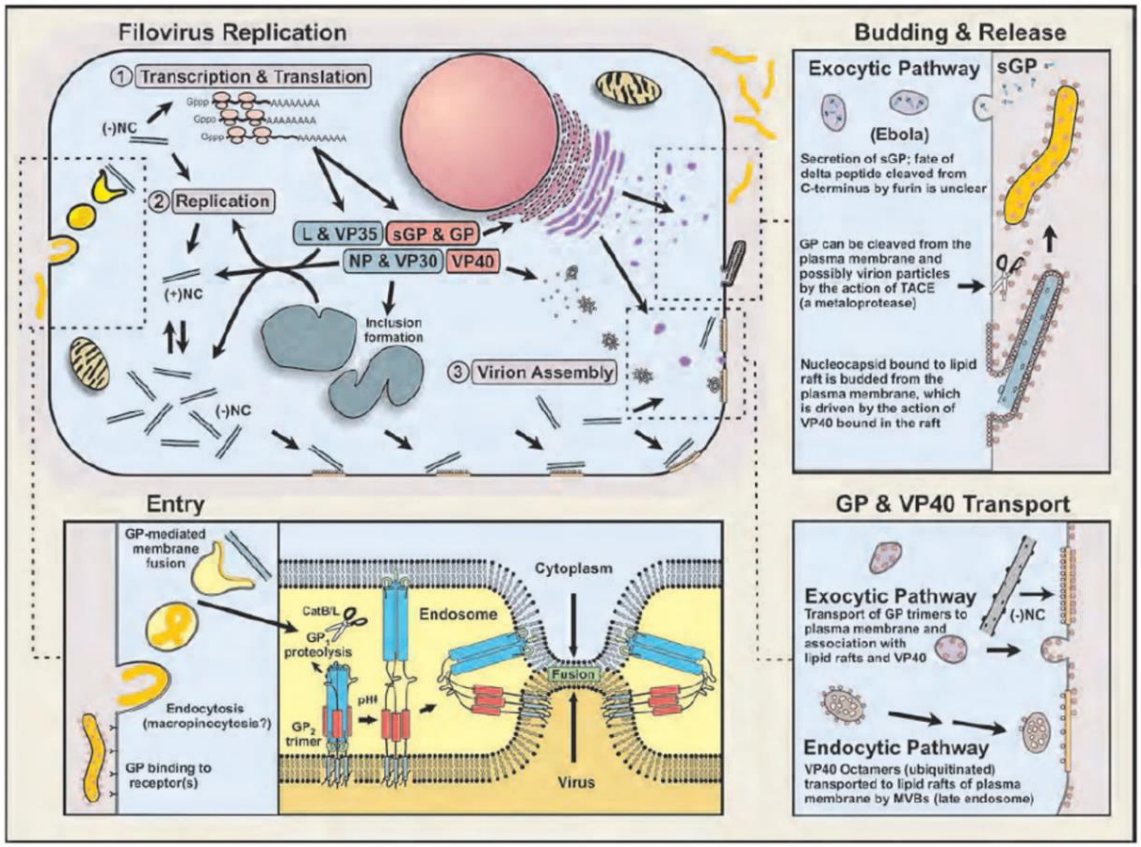


Figure 1.13: Representation of filovirus lifecycle including replication, entry, egress and specifics of GP and VP40 vesicular transport. EBOV entry is mediated by GP triggering an active transport component of the endocytic pathway called micropinocytosis. Once inside cellular vesicles GP is cleaved by furin cleavage to reveal secondary binding sites once mucin domains are removed to interact with the vesicular NPC1, and permit escape into the cellular cytoplasm. Once in the cytoplasm viral proteins are produced and processed through the Golgi and inclusion bodies form, major sites of viral replication and production of progeny virus. Viral particles escape the cell via the exocytic pathway mediated by metalloproteases. The virus buds from the cell, leaving a lipid raft on the surface of the cell. The virus carries with it phosphatidylserine, a phospholipid usually only found on the surface of apoptotic cells, or sequestered on the inner cellular membrane. This image was adapted from Fields Virology, 6th Edition (2013) (108)

1.4 Pathology

1.4.1 Pathogenesis

Typically EBOV enters a susceptible host via mucous membranes or abrasions in the skin (45), therefore EBOV has evolved tropism for cells found in peripheral tissues (109). EBOV displays high tropism for macrophages and dendritic cells (DC) early in infection (109, 110). Other cell types such as lymphocytes and unstimulated circulating monocytes which are non-adherent (111, 112) are entirely refractory to infection. The combined effects of macrophage and dendritic cell infection result in a dangerous multi-system illness (109), where patients may suffer severe haemorrhagic manifestations including mucosal bleeding and petechial rash (12, 113, 114). Patient fatality is believed due to hypovolaemic shock or multiple organ failure (24).

1.4.2 Dendritic Cells

DC infected with EBOV are prevented from maturing (115, 116). DC migrate, aiding in viral dissemination and move to lymph nodes, where pathogenic signalling results in bystander lymphocyte death (117). DC inflammatory signalling is undetectable (109, 110) the interferon response is suppressed (63, 101), and co-stimulatory proteins for lymphocyte activation are inhibited (109), this is believed to contribute to the observed massive lymphocyte apoptosis. This effect is summarised in figure 1.13 (116).

1.4.3 Macrophages

In contrast with dendritic cell infection, macrophages mature (109, 112). Activated macrophages in EBOV infection are known to hyper-secrete inflammatory cytokines including TNFα, IL6, IL8 and MIP1α (116). Infected macrophages also undergo apoptosis (109, 118, 119) late in infection, and are unable to secrete interferon (82, 109, 120) or respond to it (63, 121).

The role of macrophages in the vascular leakage has been hypothesised. Macrophage secretion of TNFα is believed to activate endothelial cells, weakening endothelial junctions (109, 122). Vascular leakage due to dysregulation of endothelial cell activation has been described in EBOV infection, facilitated by the combined effect of virally induced production of TNFα and production of sGP (122, 123).

Macrophages are also believed to have a role in coagulopathy seen in EVD (116). Macrophages are potent producers of coagulation factors including stimulators of Tissue Factor (TF) production (124-126). Production of TF and inflammatory cytokines, in addition to liver pathology is believed to result in disseminated intravascular coagulation, as infected cells migrate around the body, leading to organ failure and death (24, 125-127).

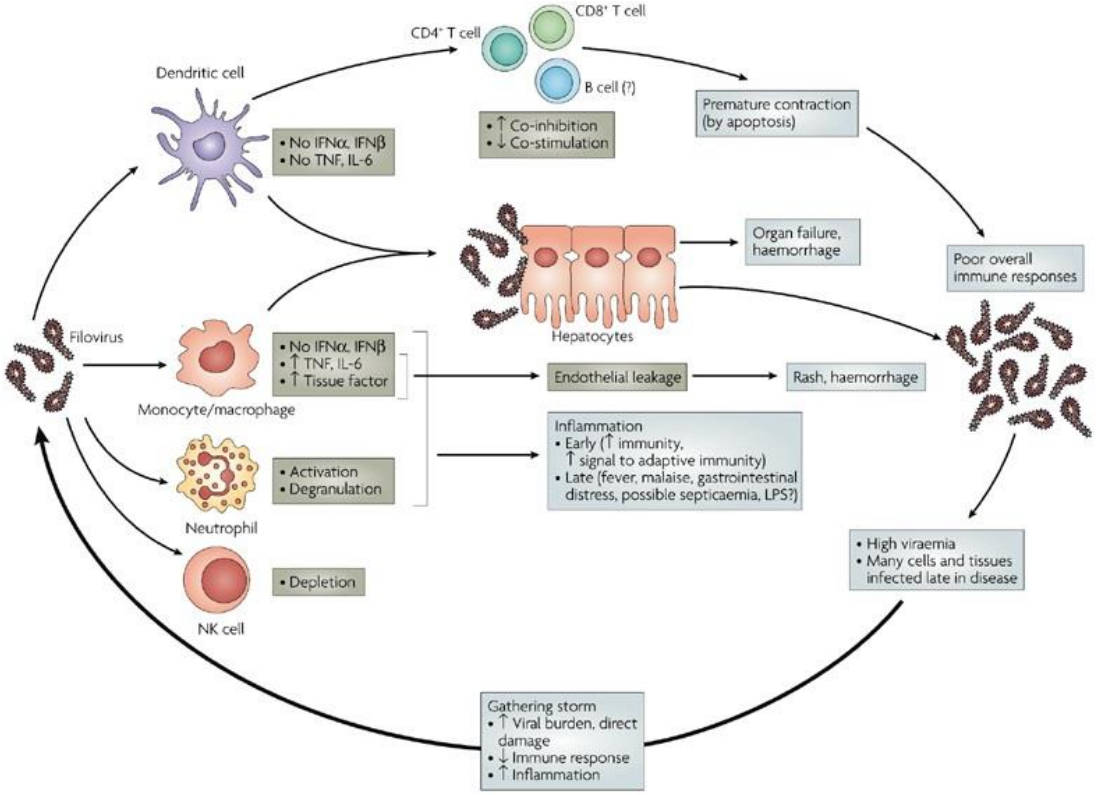


Figure 1.14: Schematic diagram showing the role of macrophages and dendritic cells in ebolavirus infection. EBOV shows high tropism for macrophages (though not monocytes), and dendritic cells. Lymphoid cells are stimulated directly by the virus or by APC, but are not direct cell targets for infection. Infection of dendritic cells leads to inhibition of maturation, no pro-inflammatory cytokines are produced nor an effective antiviral response. Dendritic cells fail to stimulate an effective lymphocyte response, leading to premature contraction of lymphoid populations and a reduction in antiviral response in the host. Infection of macrophages on the other hand leads to systemic migration, viral dissemination and production of coagulation, inflammatory and chemotactic factors. Pro-inflammatory cytokines lead to activation of vascular endothelium and vascular leakage. Inhibition of coagulation cascades and prothrombin activation leads to coagulopathy. Systemic dissemination of virus causes a rapid and diffuse inflammatory reaction, causing organ damage and haemorrhagic fever. Ineffective immune response and reduction of lymphocyte populations leads to high viral load, which is believed to correlate with poor outcome in human patients. This figure was adapted from Mohamadzadeh et al (2007) (116).

1.4.4 Inflammatory Response

EVD is characterised by systemic release of inflammatory cytokines and chemokines (6, 24). IL6 is a potent modulator of the acute phase response (128) and hypersecretion has been associated with fatal human infection with EBOV (114, 124). In addition to IL6; IL1β, IL2, IL8 (CXCL8), IL10, MIP1α, and IP-10 (124) have been associated with severe disease. Additionally, high levels of type I interferons have been shown in human patients with EVD (6, 12, 113, 114, 129). This observed hypersecretion of inflammatory activators has led to the concept of a "cytokine storm" like illness (24, 130) such as is seen in severe influenza infection (131). Parallels have also been drawn with sepsis (132), particular when combined with the full picture of vascular leakage and coagulopathies. Sepsis is a severe multi-system disease with rapid onset, due to a systemic infection with a pathogen (133), with similar symptoms to those described in systemic and severe viral infection. These parallels may lead to the application of methods for treating sepsis to the supportive therapy used for EVD patients.

1.5 Molecular pathology

Many studies have attempted to define the molecular changes in hosts infected with EBOV. Overtly, the result of EBOV infection of cultured human cells is cell death (119, 134) causing visible cytopathic effect in cell culture figure 1.14. However, early in infection EBOV appears to avoid activating the apoptotic response in cell infections (134). This is an advantage, and EBOV suppresses the activation of interferon by intracellular RNA virus detection (101), preventing cellular response to interferon signalling from other cells (63, 81).

A strong inflammatory response is a hallmark of EBOV infection, and this is believed due to signalling via cellular toll-like receptors (135), which follows similar patterns to those observed in bacterial septicaemia induced by LPS (133). The physiological changes induced by other *ebolavirus*es have not been well studied.

1.5.1 Transcriptomics

Transcriptomics involves the use of next-generation sequencing or microarrays to calculate the abundance of all transcripts encoding a gene product in a cell, tissue or organism. A greater abundance of transcript, suggested reduced gene repression, or may suggest enhanced gene transcription. Therefore the activity of pathways and transcription factors regulating a gene can be determined. Transcriptomics have been used heavily in the last decade to describe the host response to *ebolavirus* infection in both cell culture and *in vivo*. Most studies have relied on the use of microarrays (118, 136, 137). This method has many limitations, greatest of which is that the method provides only a relative level of transcript abundance compared with a control, and strict cut-off values must be used to distinguish signal from background (138, 139). This problem also occurs in polymerase chain reaction based methods, and transcripts with high abundance which may change by a large proportion relative to other transcripts may, in relative terms, fall within background level changes. RNASeq removes these barriers by quantifying molecules of transcripts produced proportionally through a PCR reaction, rather than relative abundance assessed using signal intensity (140, 141).

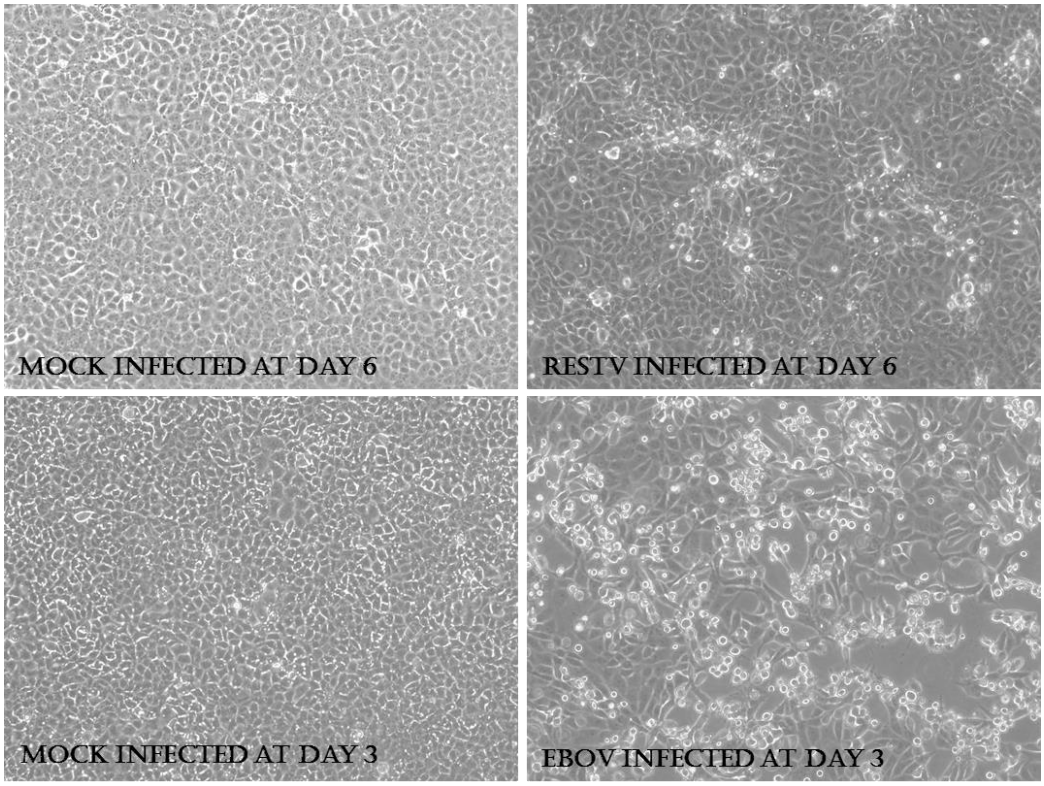


Figure 1.15: Images of Vero cells infected with either EBOV-Kikwit or RESTV-Pennsylvania for 6 days. Images were captured on days where cytopathic effect (CPE) had become noticeable. RESTV CPE is easily visible by day 6 while EBOV infected Vero cells suffer CPE after only 3 days. Images of mock infected Vero cultures are provided for interpretation at both 3 days and 6 days post mock infection. Images courtesy of Dr. Verena Krähling, University of Marburg an der Lahn, Germany, provided July 2015. Permissions are shown in Appendix I.

This means it can be determined whether a transcript with 1 million reads increased to 1.5 million reads, and may be considered a significant change. Whereas, if absolute quantities are not known, a relative increase of 1.5 fold change (FC) would be seen, below most thresholds in microarray experiments. Microarray data is also marred by inconsistencies, and experiments have sometimes shown contradictory results.

1.5.1.1 in vitro

Studies *in vitro* using transcriptomic methods have primarily focussed on EBOV, with only limited host response data on other *ebolaviruses*. Studies have failed to show large differences in the host response to EBOV compared with RESTV (136). Comparison of the host response to EBOV and MARV in human cell lines shows similar patterns of transcript abundance amongst interferon, coagulation and inflammation associated genes (115). A detailed analysis of EBOV was conducted in human peripheral blood mononuclear cells (PBMC), where the VP35 interferon antagonising domain was mutated to remove function (121). This mutant virus showed a defective ability to suppress interferon, and RNASeq data provided detailed evidence that this was the case. This analysis highlighted the high levels of inflammatory chemokines and cytokines produced during EBOV infection, and mirrored that observed in clinical EVD (124).

Thus some cell culture model experiments have provided insight which may be clinically relevant. At the outset of this project an analysis of the Makona variant of EBOV was lacking, as was a contrast between different EBOV variants, or the application of RNASeq to study RESTV infection.

1.5.1.2 in vivo

Applying transcriptomic methods to study *in vivo filovirus* infections has highlighted the diversity of diseases in animal models. RNASeq analysis of primates (142), infected with un-adapted filoviruses, guinea pigs infected with adapted strains (69, 104) and mouse models infected with adapted EBOV (143) have shown inconsistent results across models. Application of transcriptomic methods have highlighted the potential involvement of many more genes in EBOV infection, compared with traditional targeted qRT-PCR approaches used previously. For example; transcripts encoding metalloproteases were over-represented in lethal EBOV infection of mice (143). The role of interferons has also been studied in the primate model, illustrating some important differences in the transcription of interferon associated genes in a living host compared with cell culture during EBOV infection (137, 142). Experimental evidence has highlighted a number of proteins which are host interacting partners with EBOV proteins, and are detailed in figure 1.15.

1.5.2 Proteomics

Proteomics has been used to validate transcriptomic data modelling infection of cells with Hendra virus (151), and have been used to reveal novel interactions identified by finding unusual patterns of post-transcriptional modulation (152, 153). Proteomics applied to filovirus research is limited. The application of proteomics to study host-virus interactors in conjunction with co-immunoprecipitation has identified important interactions of VP24 with Karyopherin-α, (154) and interactions of NP and VP24 with heat shock proteins (64, 102, 155). The role of chaperones in viral infection is not unexpected, as they fulfil a physiological role in stabilisation of proteins, and viral protein stability is evolutionarily advantageous.

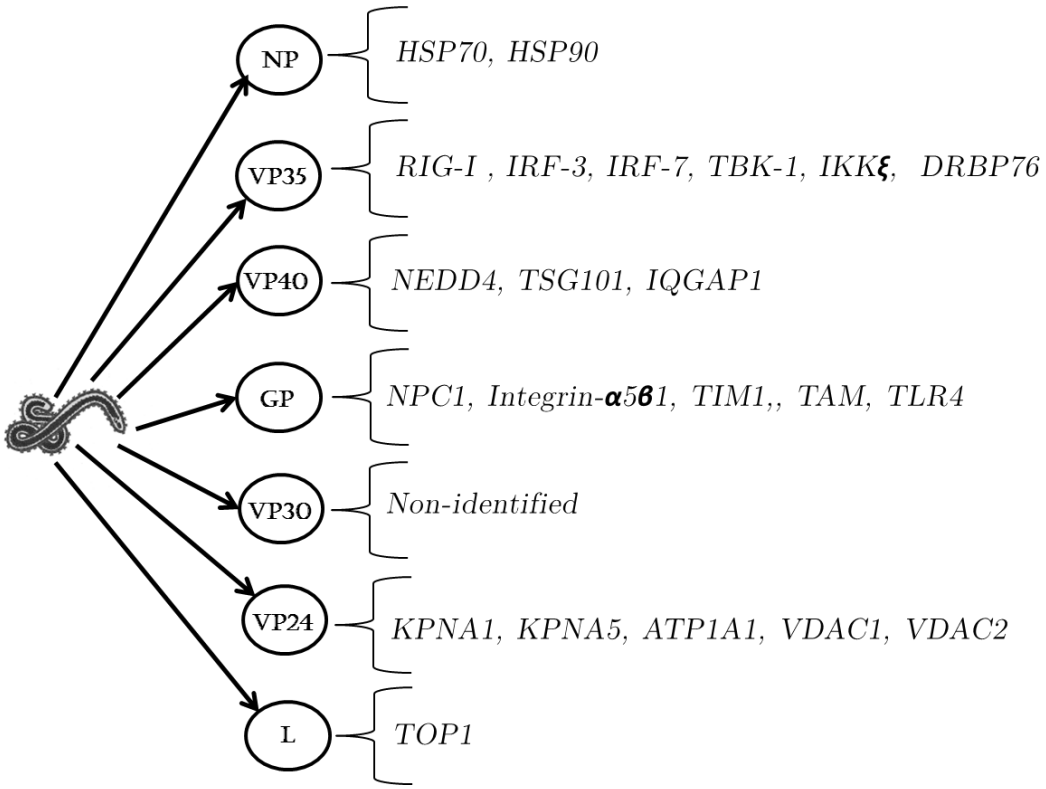


Figure 1.16: Schematic diagram illustrating the ebolavirus proteins and host proteins with which they interact. Several studies have shown the importance of host factors in ebolavirus lifecycle, not least of which are the interactions of VP24 and VP35 which suppress the antiviral response in cells. Shown are the common gene identifiers for human proteins on the right, next to each ebolavirus protein. In literature searches no published studies on the interactors for EBOV VP30 could be found. (64, 85, 95, 97, 99, 102, 135, 144-150)

A high depth proteomic analysis of viral particles revealed a complex array of host proteins are incorporated into virion membranes or packaged within, as they egress and infect new cells (156). Some of these proteins fulfil the role of surface receptors, and as such may impart an enhanced ability to enter cells later in infectious lifecycle, and should be studied further. HSPA5 and RPL18 were shown to have relevant virion incorporation, as their inhibition resulted in antiviral effects (155, 156). There are no studies which report the application of proteomics to *in vivo* infection with *ebolaviruses*; likewise there are no proteomic studies which have been applied to study other *ebolaviruses* such as RESTV in cell culture.

1.5.3 Interferons

Interferon signalling is the principal antiviral defence mechanism in mammalian cells. Interferon signalling is subdivided into type I (principally IFNα/β), type II (IFNγ) and type III (IL-28A, IL-28B and IL-29) (157). The production of dsRNA by both DNA and RNA viruses during viral replication and transcription is a potent inducer of the type I interferon response (158). dsRNA sensing is mediated by MDA-5 or RIG-I intracellular receptors, which contain a CARD, recruiting the Cardif/MAVS/VISA/IPS-1 complex (157). The complex then mediates the transcription of downstream response genes by IRF-3, IRF-7 or NFκB (157).

MyD88 is a major inducer of NFκB and IFNβ production via IRF-3/7 signalling (159). MyD88 is recruited to activated TLR7 and TLR9 found on the inner surface of endosomes and activation induces NFκB and IRF7 signalling events (157). TLR3 is also found on the interior surface of endosomes, and when activated signals to TRIF, which mediates downstream NFκB and IRF3/7 activity. These interferon activation processes are

summarised in figure 1.16. EBOV is able to suppress interferon through the antagonistic function of two proteins: VP24 prevents STAT1/STAT2 nuclear translocation and VP35 interrupts RNA sensing via RIG-I and also prevents regulatory function of IRF-3 (63, 101).

A study comparing the crystal structures of EBOV and RESTV VP35 interferon inhibitory domain showed near complete homology (160), and studies of EBOV and RESTV VP24 and 35 showed very similar sequences (25, 105). However, studies have suggested that there may be a slight decrease in effective interferon antagonism in RESTV compared with EBOV (160). These slight differences may contribute to the dichotomy in pathology observed in humans infected with these two viruses.

The study of interferon signalling in human patients, animal models and in cell culture has shown that in living hosts interferon signalling is highly active (12, 16, 113, 114, 129, 132), in contrast to the near total suppression seen in cell culture. Interpretation of this difference is that infected cells suffer sustained interferon inhibition, while bystander cells may be activated by other factors induced by cell death and inflammatory signalling. Interferon suppression is possibly linked with activation of an NFκB regulated inflammatory response (135, 161-164), which is pro-cell survival (165), and thus may counteract the apoptotic induction by interferons, and contribute explain the lack of apoptosis in EBOV infected cells (119, 134, 166). Much transcriptomic work has focussed on interferon signalling, indicating similar patterns of gene transcription in EBOV, MARV and RESTV (101). Further study of interferon signalling in the cells of humans and other host species susceptible to RESTV infection may highlight differences in viral fitness, which may help identify species specific factors.

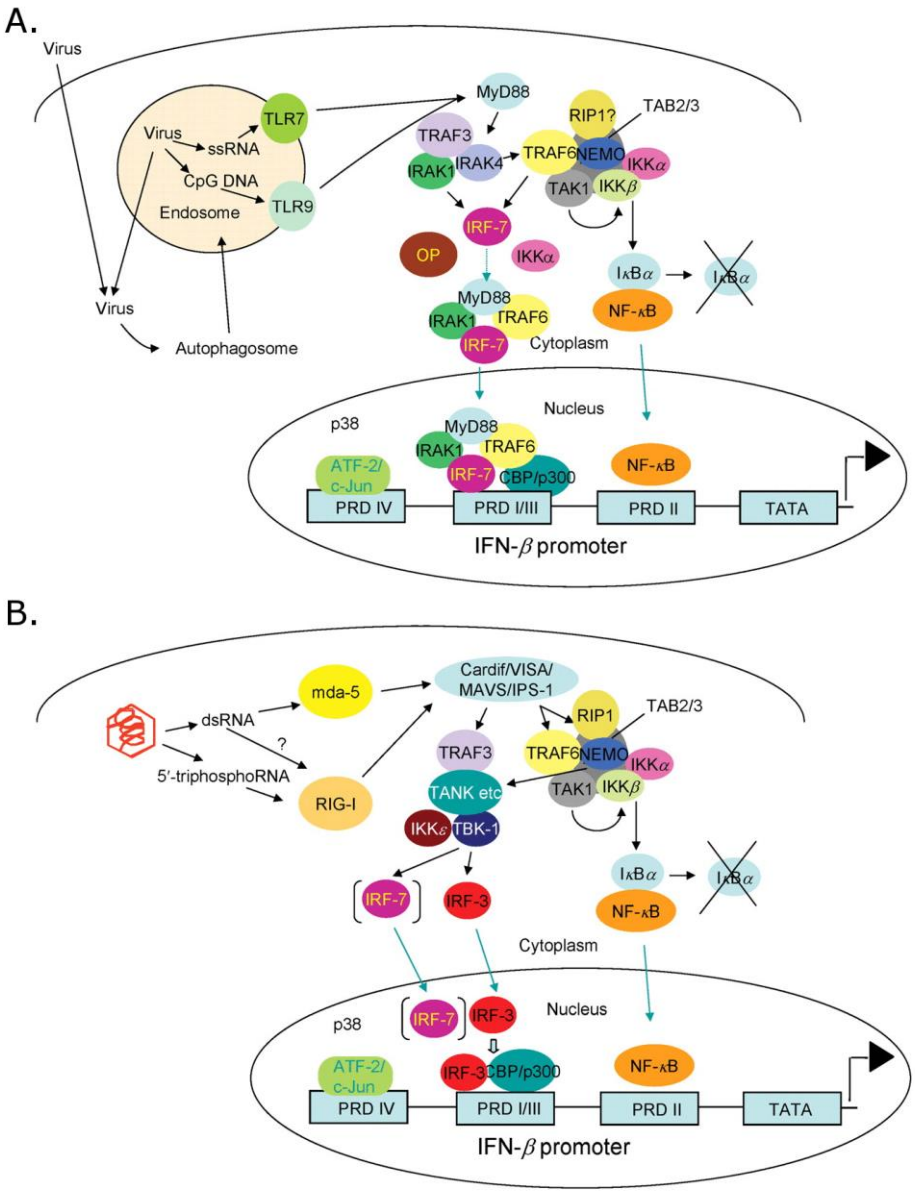


Figure 1.17: Schematic diagram of type I interferon signalling. (A) Signalling involving processes where the virus uncoats and replicates inside the endosome, producing ssRNA or CpG DNA which activate the sensing receptors TLR7 or TLR9 on the interior surface of endosomes. Activated TLRs recruit MyD88 which facilitates downstream signalling to NFκB and complexes with IRF7 to transcribe from the interferon-β promoter, and activate the type I interferon response. (B) A depiction of events seen in cellular infection with EBOV and other RNA viruses. Viral uncoating and replication/transcription results in detection by mda-5 (dsRNA) or by detection of t-triphosphateRNA by RIG-I. Activated RNA sensing receptors recruit the Cardif/VISA/MAVS/IPS-1 complex to a CARD domain, which facilitates downstream signalling to NFκB, IRF3 and IRF7. This figure was adapted from Randall & Goodbourn 2008 Published by the Microbiology Society, UK (157).

1.5.4 Immune Response

In addition to interferon signalling events, several genes regulate the immune response to infection or stress in cells: NFκB, AP-1, SP-1, PKR and EIF2 (157, 165). The stress response has been connected to viral infection as a possible antiviral mechanism (167, 168) as well as a possible virally induced method of limiting host protein production (153, 169). Stress granule formation has been evolved to limit the abundance of host and virus encoded proteins by sequestering untranslated messenger RNA complexes in stress granules (170). An early signalling event in stress granule formation is the phosphorylation of EIF2α by the RNA activated protein kinase PKR. Phosphorylated EIF2α then interrupts the activity of EIF2B (guanine nuclear exchange factor) which forms the ternary complex (167, 168, 170). Interruption of the ternary complex leads to a reduction of normal cellular protein production, and a bias to production of host stress response factors. The interaction of heat shock proteins with EBOV proteins (64, 155) may indicate that the stress response is required to enhance viral protein stability. There has not been a detailed study of the proteomics of EBOV viral infection, and proteomic analysis of EBOV infected cells would thus answer whether the stress response is a key response factor in infection, as it might be expected host protein synthesis would decrease. Prevention of the antiviral response is essential for successful viral infection. TNFα signalling is a prominent feature of clinical EVD (122, 124, 129). AP-1 may be activated via the DDX58 (RIG-I) signalling pathway, as well as by interaction of TNFα, TGFβ and IL2 with cellular receptors, which also leads to activation of NFκB and interferon-α/β signalling (165, 166). NFκB activation is likely to significantly impact EBOV lifecycle (165).

1.5.5 NFκB signalling pathway

The NFκB complex is named nuclear factor kappa-light-chain-enhancer of activated B cells. NFκB exists as a complex of several components, the final configuration of which determines the regulatory effects on transcription. All proteins of the complexes share a Rel homology domain (RHD) in the N-terminal region (165). This domain is shared by other mammalian transcription factors such as NFAT, and phosphorylation of this region controls transcriptional activity (165, 171).

NFκB components form a family of proteins which include RelA (p65), RelB and c-Rel (encoded by RELA, RELB and REL respectively). A further two components, NFκB1 and NFκB2 (NFKB1 and NFKB2) are synthesised as larger proteins called p105 and p100, these are then processed via the ubiquitin/proteasome pathway into p50 and p52, which act as transcriptional repressors when bound to RHD containing components (RelA, RelB, c-Rel) (165). Complex formation is constitutive but may be enhanced by phosphorylation during NFκB pathway activation (172). The canonical activation route for NFκB complexes is via interaction of TNF or IL1 with surface receptors, or activation of TLR4. An alternative activation pathway exists triggered by interaction of CD40L with CD40 on the cell surface (173).

Migration of the NFκB complex into the nucleus to commence transcriptional activity is constitutively suppressed by interaction with IκB. IκB is a family of proteins which must undergo ubiquitination and degradation prior to NFκB transcriptional activation. Inhibitory activity is mediated by Ankyrin repeat domains located in the C-terminal region of IκB proteins. Interestingly these Ankyrin repeat domains are also present in unprocessed p100 and p105, granting similar inhibitory effects on NFκB complexes (165).

Canonical NFκB activation results in degradation of IκB proteins and unbinding of the master regulatory protein NEMO (IκBγ). This allows heterodimers composed of RelA (p65)-p50 to migration into the nucleus. This configuration results in transcription of genes involved in inflammation, immune response and cell survival. Conversely, activation of the alternative pathway results in cleavage of p100 into p52, and formation of RelB-p52 complexes promoting the transcription of genes involved in the adaptive immune response. These processes are shown in a schematic diagram in figure 1.17 (165). Dimers composed of p50-p50 have been shown to repress activation induced by IL1β. P50 dimers therefore may serve as a regulatory mechanism to control p65-p50 transcriptional activity and modulate the immune response. The NFκB complex controls transcription of >100 genes. Control of NFκB signalling is therefore a distinct advantage in viral infection (174).

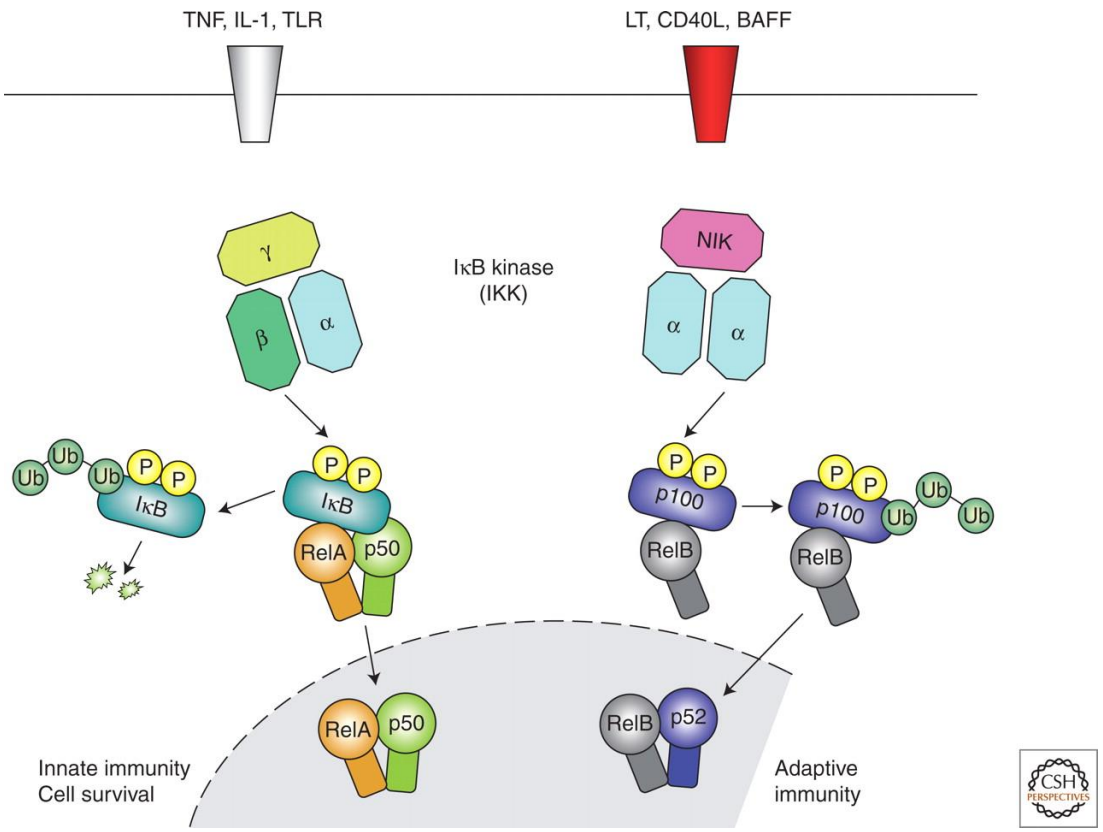


Figure 1.18: Depiction of the canonical (left) and alternative (right) NFκB signalling pathways. Activation of the canonical pathway takes place through interaction of ligands with TLR (such as TLR4) or receptors for TNFα or IL1 on the cell surface, resulting in the activation of RelA (p65) which mediates activation of response genes involved in immune response and promotion of cell survival. In contrast the alternative pathway is activated by interaction of cell surface receptors with CD40L (CD40), resulting in the activation of genes involved in B-cell signalling from antigen presenting cells and the regulation of the adaptive immune response. This signalling is critical in processes involving T and B-cell selection in immune stimulation by dendritic cells. This schematic was adapted from Lawrence (2009); Cold Spring Harbour Publications (165)

1.5.6 Toll-like receptors

TLR (toll-like receptors) are the primary pathogen associated pattern recognition receptors in mammals (175). The role of TLR in *ebolavirus* infection was highlighted in studies of EBOV GP, showing association of GP and sGP with TLR4 (135). These studies also showed that sGP was required for cell death (94). sGP is also a necessary factor for pathogenicity in primates (176), but not in guinea pigs (177), and is likely also needed for pathogenicity in humans. Reducing sGP production by mutations in the EBOV editing site resulted in attenuation. The high expression of TLR4 on the surfaces of macrophages and dendritic cells (primary cell targets after host entry) means the role of this receptor in EBOV infection and lifecycle is an interesting target for further study (135).

TLR4 signalling also results in NFκB activation (135, 174), resulting in enhancement of cell survival and production of highly potent inflammatory signalling cytokines including IL-6 and TNFα, the roles of which have long been known for EBOV infection as keystones of the immunopathology (109, 124). Studying the impact of *ebolavirus*es on these immune-pathogenic pathways may permit hypotheses to be generated as to the expected pathology seen in host species.

1.6 Research project

This research project was carried out to characterise the host response to the Makona variant of EBOV, developing a greater understanding of the pathogenesis of *ebolaviruses,* making use of high resolution molecular methods, clinical cases and well characterised host cell models, to elucidate the molecular mechanisms underpinning *ebolavirus* infection.

1.6.1 Hypothesis

I. The host response to the Makona variant of Ebola virus from West Africa was distinct from previously isolated *ebolaviruses*

1.6.2 Project Aims

I. Identify differences in the host response to Makona variant of EBOV and other *ebolavirus*es through high resolution molecular analysis

1.6.3 Project Objectives

I. Profile host transcriptomic and proteomic changes during infection with the Ebola virus, Makona variant.

II. Contrast the host response to the Ebola virus, Makona variant and other *ebolavirus*es to identify signatures of infection unique to the Makona variant.

III. Using transcriptomic and proteomic methods, identify host factors with important roles in the lifecycle of the Ebola virus, Makona variant.

2. Methods

2.1 *In vitro* tissue culture techniques

2.1.1 Routine culture of adherent cell lines

Adherent cell lines used in experiments throughout this project were acquired from the ECACC; A549 (86012804), Vero E6 (85020206), TB1-lu (90020805), ESK-4(93120821), HEPG2 (85011430). PK-15 (ATCC CCL-3) cells were acquired from ATCC. BSRT7 cells are derived from BHK cells, stably transfected with a plasmid encoding a T7 polymerase under the power of a CAG promoter, and were kindly provided by Dr. J. Barr of Leeds University. A549 NPro cells are derived from A549 cells as above, which constitutively express the N protein of BVDV, and were kindly donated by Professor S. Goodbourn of St. Georges, University of London. All adherent cell lines were cultured in DMEM containing GlutaMAX (Thermo Scientific) supplemented with heat-inactivated FCS (Thermo Scientific) at either 10% (A549, A549 NPro, Vero E6, TB1-lu, ESK-4, HEPG2, PK-15) or 5% (BSRT7) concentration. No antibiotics were used in routine cell cultures however BSRT7 cell culture media was supplemented with 1:1000 dilution of G418 Gentamicin derivative (Thermo Scientific) selection agent, to select for plasmid expression of resistance genes. Incubation was performed in a serviced humidified incubator (Thermo Scientific) maintained at 5% CO_2 and 37°C. Cell detachment for passage was mediated by 0.1% Trypsin-EDTA premixed reagent (Thermo Scientific) incubated for between 3-10 minutes until full detachment of the cell monolayer and cells were counted using plastic disposable C-Chip Neubauer cytometers (NanoEnTek).

2.1.2 Routine culture of semi-adherent/suspension cell lines

Semi-adherent/suspension cells used in this project included THP-1 (88081201). THP-1 cells were acquired from ECACC and an authenticated, mycoplasma and bacteria negative cell bank was set up for this project by ECACC technicians. All cultures were started from this authenticated stock to ensure standardisation of experiments across the project. For passaging, cells were reseeded at a density of 0.1-0.3x10e+6 cells per ml. Cultures were maintained between 0.3x10e+6 and 2x10e+6 in RPMI media supplemented with GlutaMAX (Thermo Scientific) and FCS (Thermo Scientific) at 10% for routine culture or 20% at initial resuscitation from frozen ampoules.

2.1.3 Generation of PMA-differentiated THP-1 macrophage-like cells

Naïve THP-1 cells are semi-adherent cells morphologically and functionally similar to Monocytes (178-180), and can be differentiated into macrophage like cells using Phorbol-12-myristate (PMA) with optimised protocols (179). Lyophilised PMA powder (Sigma-Aldrich) was reconstituted and diluted in DMSO to a working concentration of 1µg/ml. PMA was diluted to a final concentration of 5ng/ml in RPMI supplemented with 10% media. Cells were suspended in complete growth media with additive 5ng/ml PMA to a concentration appropriate to intended seeding density. After thorough mixing by inversion cells were added to wells/flasks to form 80% to 90% confluent layer of cells. PMA differentiated THP-1 cells do not form a pavement monolayer and after activation with PMA will suffer a severe inhibition to growth rates. 24h after addition of PMA, cells were rinsed twice with PBS (Thermo Scientific) and fresh complete RPMI growth media added. Cells were rested for 72-96h before treatment, transfection or infection took place. During this resting period the cells slowly differentiate, changing morphologically and becoming adherent, this change is shown in Figure 2.1.

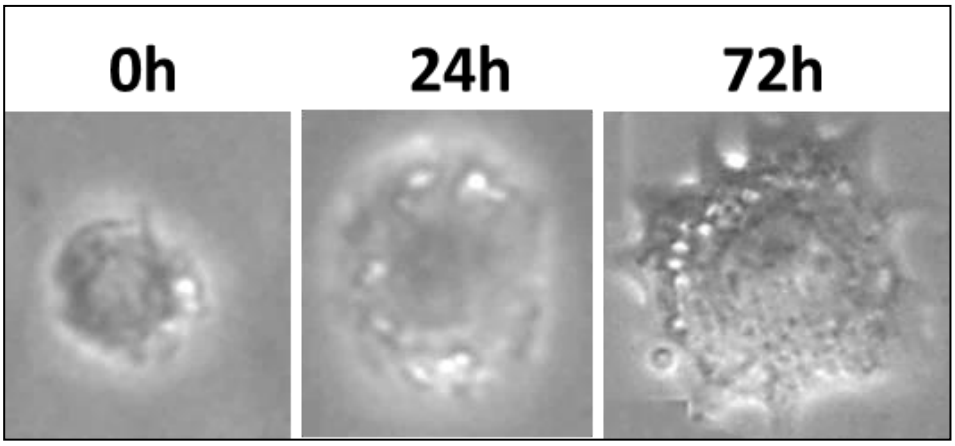


Figure 2.1: Example of THP-1 differentiation and changing morphology after exposure to 5ng/ml phorobol-12-myristate for 16-24h followed by 72h rest period. 0h indicated an image captured immediately following rinsing of the now adherent cell layer with PBS and addition of fresh complete growth media, at this stage cells are now adherent, but still exhibit small-round cellular morphology as seen in unstimulated THP-1 cells. A second image was taken 24h after commencement of the rest period, indicating slight changes in cellular morphology. By 72h the cells have fully differentiated and become fully adherent. At this stage it was difficult to dislodge cells from plastic without mechanical disruption and aggressive treatment. The cells are larger, clump together and have projections indicative of adherence.

2.1.4 Cell culture at Containment Level 4

Cell culture at CL4 was complicated by the severe pressure differential often present inside the fully contained cabinet line system, alongside high concentrations of formaldehyde residues and cytotoxic chemicals. The result was that where possible, cell growth should be in flasks with plug-sealed caps rather than filter caps, and cultures should be locked in sealable boxes. Where plates have been used in this project, plates were sealed with tape and placed in re-sealable bags before placing in o-ringed boxes. With these precautions cell viability in negative control cells remained >90% throughout culture periods. Cell viability was evaluated using Trypan Blue (Sigma-Aldrich) and manual counting using a cytometer.

2.1.5 Viral Infection

The EBOV Makona variant was obtained from the European Mobile Laboratory and was isolated during the laboratory response to the 2014 West African outbreak through work described in section 3.1. This isolate was designated Ebola virus/H. sapiens-wt/GIN/2014/Makona-Gueckedou-C05. EBOV-Ecran was isolated during an outbreak in October 1976 (WHO International Commission, 1978) designated Ebola virus/H.sapiens-tc/COD/1976/Yambuku-Ecran. Reston virus Pennsylvania variant M. fascicularis-wt/USA/1989 /Pennsylvania was acquired from the Centre for Disease Control, Atlanta, USA. All viruses have variable passage history but have all been cultured only in Vero cells. Stocks were generated in Vero cells for all viruses used in this project, stocks were mycoplasma tested and authenticated by sequence analysis before use, and used as a standard stock throughout the project. Viruses were quantified by a modified $TCID_{50}$ method discussed in section 2.1.6.

Viral inoculums were prepared in ACDP CL4 in serum free DMEM supplemented with 3% HEPES and inoculated onto cell cultures, cells were prepared by washing with PBS and transferring to ACDP containment level 4, then PBS replaced with viral inoculum and incubated for 60 minutes at 37°C. After 60 minutes complete DMEM or RPMI growth media was added as directed in section 2.1.1 and 2.1.2.

2.1.6 $TCID_{50}$

A modified $TCID_{50}$ assay protocol was employed at ACDP CL4 to avoid the use of 96 well plates, which generally suffer viability issues due to the toxic environment. Assays were conducted in plug-sealed capped flasks with a 6 point dilution series, each with 5 replicates. Viral titres were calculated using $TCID_{50}$ calculations performed using the Reed & Muench method (181).

2.1.7 Culture on multi-well chamber microscopy slides

To culture adherent cell lines and differentiated THP-1 cells on a glass slide for subsequent microscopy and fluorescent assays, cells were seeded and cultured according to conditions specified in sections 2.1.1, 2.2.2 and 2.1.3. 8-well chamber slides were used for these experiments (Nunc, Thermo Scientific), specifically Lab-Tek II 8-well chamber slides (154534). These slides are fixed to a silicon gasket which forms the culture wells, onto which a plastic lid was placed to allow aseptic tissue culture. Cells were seeded at ~80,000 cells/well. For culture at CL4, slides were sealed inside plastic o-ringed boxes and placed in a 37°C without CO_2.

2.1.8 Fixed cell staining for viral proteins

A549, PK-15, ESK-4, TB1-lu or differentiated THP-1 cells were seeded into each well of an 8 well chamber slide (Ibidi, UK) at a density of 20,000 cells/well in complete growth media as described in 2.1.1 and 2.1.2.

Slides were incubated for 16h in a containment level 2 laboratory before relocation to containment level 4 and infected with EBOV or RESTV as described in 2.1.5. After 24h post infection, slides were washed twice with sterile PBS (Thermo Scientific) and submersed in 10% Neutral Buffered Formalin (NBF) (Sigma-Aldrich), a concentration of minimum 4% formaldehyde: paraformaldehyde and methanol.

Slides were fixed at 4-8°C for a minimum of 72h in accordance with locally agreed risk assessment, but for a maximum for 28 days. Slides were then removed from 10% NBF, rinsed with sterile PBS and submersed in Acetone for 16h, before removal, drying and freezing until required for use.

Cells were stained for EBOV or RESTV GP using the Leica BondMax (Leica Biosystems) and the Leica Bond Polymer Refine Detection kit (Leica Biosystems). An antigen retrieval step was included for 10 min using the Bond Enzyme Pre-treatment kit. A rabbit polyclonal, anti-EBOV or anti-RESTV GP antibody (IBT Bioservices) (dilution 1:1000) shown in Table 2.3 was incubated with the slides for 60 min. 3,39- Diaminobenzidine chromogen (dark red) and haematoxylin (pink) counterstains were used to visualize the slides.

2.1.9 Use of reporter cell lines

HEK-293T cells stably transfected with a lentivirus encoding a secreted embryonic alkaline phosphatase enzyme (SEAP) downstream of either the IFI53 promoter driven by Interferon α or Interferon β (IFNα/β), or the under the control of NFκB and AP-1 transcription factors. These HEK-Blue cells (Invivogen) were used to measure the levels of bioactive TNFα and IFNα/β present in secretions. HEK-Blue cells were cultured in accordance with manufacturer instructions in complete DMEM growth media and 10% FCS supplemented selective antibiotics. TNFα reporter cells required 100µg/ml Zeocin (Invivogen), and 10µg/ml Pluromicin (Invivogen).

IFNα/β reporter cells required 100µg/ml Zeocin, and 30µg/ml Blastocidin (Invivogen). Culture was performed as described in 2.1.1. HEK-Blue were grown to 80% confluence, washed twice with sterile PBS and treated with 20µl supernatant collected from infected/treated/transfected cells, diluted in complete growth media. After 24h supernatant was collected from HEK-Blue, and the amounts of SEAP measured using the Quant-BLUE assay kit (Invivogen). TNFα or IFNβ2 were used to treat the relevant reporter cell line directly at a range of 5000nM, 2500nM, 1250nM, 625nM, 312nM 156nM, 78nM or Untreated, in order to determine a suitable reactivity curve as a positive and negative control for this assay.

2.2 RNA extraction and purification

2.2.1 Viral RNA Extraction

For extraction and purification of viral RNA, the QiaAMP Viral RNA Mini Kit (Qiagen) was used. This kit was validated as a suitable method for extraction of viruses by diagnostic clinical laboratories and the ability of the contained Buffer AVL to neutralise EBOV has supporting evidence in the literature. All extractions were performed to the manufacturer's specifications. 140µl cell culture supernatant or patient serum was added to 560µl Buffer AVL (Anti-Viral Lysis).

After a 10 minute incubation period at room temperature, 560µl of 100% molecular biology grade ethanol was added to the sample, and mixed by trituration. Ethanol addition causes the nucleic acid to precipitate from the sample. The treated samples were then centrifuged on the provided filter columns at 8000 xg for 30s-60s. Wash steps with buffers containing guanidinium isothiocyanate and ethanol were performed twice with Buffer AW1 and Buffer AW2 provided with the kit. Viral RNA was eluted in 60µl Buffer AVE or TE buffer.

2.2.2 Cellular RNA Extraction

Cellular RNA was extracted and purified using the RNEasy Mini Kit (Qiagen). This system utilises Buffer RLT as an extraction buffer. Buffer RLT was validated to neutralise EBOV and was able to lyse cells to extract RNA from the cellular fraction of cell culture and patient samples. Buffer RLT was supplemented with 0.02% β-mercaptoethanol (Thermo Scientific) to increase denaturing potential and improve RNA stability by further disrupting activity of any RNase enzymes. As discussed in 2.2.1 all kit instructions were followed while performing these extractions. The optional DNase treatment step was omitted from the extraction process, as this was carried out when required for qRT-PCR or RNASeq analysis. Cells were lysed with 600µl Buffer RLT added directly to cell monolayers or patient blood cell fractions. After 2-3 minutes incubation at room temperature, samples were agitated to ensure complete lysis. 10 minutes after addition of Buffer RLT, lysates were collected in a 1.5ml Eppendorf or screwcap tube, 600µl 70% molecular grade ethanol was added to the samples and the entire lysate including any precipitate was transferred to filter columns. Centrifugation steps were performed according to manufacturer's instructions, and wash steps performed with the provided buffers RPE and RW1. Cellular RNA was eluted in TE buffer and purified RNA stored at -20°C until required for downstream activity.

2.2.3 Estimating RNA concentration by Nanodrop and Qubit assays.

Cellular and Viral RNA was quantified using the Nanodrop 2000 (Thermo Scientific) or Qubit (ABi; Thermo Scientific). The standard Nanodrop 2000 assay for RNA concentration estimation was performed and ratios at 260/280nm and 260/230nm calculated to ascertain purity. Samples below an expected value of 1.8 were retested and if persistently below this level, were re-extracted; no samples used in this project failed this basic quality check.

2.3 RT-PCR techniques

2.3.1 Generation of a standard curve of synthetic RNA

For the quantitation of viral genome copies, a synthetic control of known quantity was used to calibrate and assess the results of qRT-PCR assays targeting viral genome sequences. A synthetic control was designed mimicking the region targeted by the qRT-PCR assay used in this project (182). This synthetic oligonucleotide sequence was cloned into the pMX-01 vector under the power of a T7 promoter (GENEArt; Thermo Scientific) and sequence verified by Sanger sequencing at GENEArt (Thermo Scientific) and Public Health England's Sequencing Service, Colindale, London, U.K. PCR was performed to amplify the region of interest and samples added to E-Gel EX 2% Agarose electrophoresis gels (Invitrogen; Thermo Scientific).

Samples were electrophoresed for 30 minutes at 110V in TBE buffer. Bands were visualised on trans-illuminator and excised by scalpel. Gel extractions were performed with the Gel Extraction and Purification Kit (Qiagen) to manufacturer's instructions. To produce the synthetic RNA control, the MegaShortScript T7 kit was used (Ambion; Thermo Scientific). A mastermix of the T7 reaction was made according to manufacturer's instructions and T7 "run-off" reactions performed for 4h at 37°C. 15 minutes before the end of the run off period, 2µl TURBO DNase (Ambion; Thermo Scientific) was added to the sample mix, to remove all DNA from the sample). RNA oligonucleotides were purified using the RNEasy MinElute kit (Qiagen) to manufacturer's instructions. Samples were quantified using the Nanodrop 2000 (Thermo Scientific) and concentration standardised to 2 x 10^{10} copies/µl. All work with these high concentration samples was performed in laboratories where other work with EBOV will not be performed to reduce PCR contamination risk.

2.3.2 One step qRT-PCR for viral GP

Quantitative real time reverse transcription polymerase chain reaction (qRT-PCR) was used to quantify the abundance of viral genome in cells, sera and supernatant. Rather than designing an assay *de novo* an assay with established effectiveness and with validation data supporting usage was selected. This assay has been fully validated for clinical diagnostic use in the Rare & Imported Pathogens Laboratory, part of Public Health England, U.K. and has been published in a peer reviewed scientific article (182). The qRT-PCR primers were designed as shown in the paper for EBOV and RESTV, including the MGB probe design (Thermo Scientific). 900nM Forward primer, 900nM Reverse primer and 250nM Probe were added to each reaction. The TaqMAN Fast Virus kit was used for reverse transcription and PCR (Thermo Scientific), and water used in the reaction was nuclease free water (Qiagen).

Run conditions for the assay were as recommended by the TaqMAN Fast Virus Kit (Thermo Scientific), as 50°C for 10 minutes, 95°C for 2 minutes, followed by 45 cycles of 95°C for 15 seconds and 60°C for 3 seconds. Data acquisition was performed at the 60°C step using the ABi SDS version 7.0 software (Thermo Scientific), and RT-PCR performed in one step using the ABI 7500 Fast real-time PCR thermocycler (Thermo Scientific). Threshold value used throughout the study was set to 0.2 log fluorescence normalised to ROX integrated dye (Thermo Scientific). Viral quantification was performed using the standard generated in 2.3.1 using the on-board software to compare sample cycle threshold (Ct values) to those of the standard curve. This allowed calculation of copies of equivalent viral genome corresponding to the GP gene in every µl of extract.

2.3.3 Reverse transcription

Reverse transcription assays for cytokine profiling was performed with the RT2 Reverse Transcription kit (Qiagen). This kit was used to manufacturer instructions. In brief, after any initial treatment of the extracted RNA, a maximum of 8µl volume and 1µg of RNA was combined with 2µl Genomic DNA Elimination Kit, to a final volume of 10µl and incubated in the ABi 7900 Thermocycler unit (Thermo Scientific) for 10 minutes at 42°C.

During this period the reverse transcription mix was prepared as directed in the kit manual then placed on ice for a minimum of 1 minute. The sample was removed from the thermocycler and placed on ice for a minimum of 1 minute, and then 10µl of the pre-prepared reverse transcription reaction mix added to the sample on ice. The completed reverse transcription reaction was incubated in the thermocycler at 42°C for 15 minutes, and then inactivated at 95°C for 5 minutes, before 92µl of nuclease free water (Qiagen) was added to the sample. The sample was then stored on ice until required if within a few minutes of use, stored between 3-8°C in a refrigerator if used within 48 hours, or frozen below -20°C if to be used in several days' time.

2.3.4 DNA removal and reaction clean-up

Pre-treatment with TURBO DNase (Ambion, Thermo Scientific) was required where substantial DNA contamination of RNA samples was detected. Equimolar concentrations of samples were treated with 2U of TURBO DNase for 15 minutes at 37°C, before heating to 95°C for 5 minutes and cooling on ice until required. For clean-up and inactivation of TURBO DNase, the RNEasy MinElute kit was employed (Qiagen). Samples treated with DNase were added to 350µl of Buffer RLT (Qiagen), after 10 minutes in this chaotropic salt solution, 350µl of 70% molecular biology grade ethanol was added.

The mixed sample was transferred to RNEasy MinElute spin columns cooled at 2-8°C prior to use, and the manufacturer's instructions followed. Purified RNA was eluted in 14µl Nuclease Free Water (Qiagen) and concentrations calculated using a Nanodrop 2000 (Thermo Scientific) to estimate ng/µl RNA remaining in the samples. Samples were diluted to equimolar RNA concentrations prior to downstream activity.

2.3.5 qRT-PCR for cytokine mRNA

To profile the abundance of mRNA corresponding to human cytokines the RT2 Profiler Array (Qiagen) card for Antiviral Response was used (PAHS-122). RT2 profiler array experiments were conducted in accordance with manufacturer's instructions. ABI Fast 96 well plates (Life Technologies) were prefilled with lyophilised commercial assays targeting 84 different human antiviral gene mRNA. These 84 genes are listed in the manufacturer's handbook and specification guide (Qiagen) and detailed in Table 2.2. 84 targets of interest, 5 housekeeping genes, a control for genomic DNA contamination, a reverse transcription control, 3 positive controls and a negative control were included in every test. Purified RNA was processed as in sections 2.3.3 and 2.3.4 and diluted to a total volume of 102µ with nuclease free water (Qiagen). 90µl diluted RNA was combined with 1260µl RT2 SYBR Green Mastermix (Qiagen) and 1350µl RNase free water (Qiagen). The final solution was mixed by inversion and 25µl added to every well of a RT2 Profiler Array card. Plates were agitated to help suspend the lyophilised primer/probe mix and spun at maximum speed for 30-60 seconds.

Toll-Like Receptors & Chaperones	CTSB, CTSL, CTSS, TLR3, TLR7, TLR8, TLR9 CHUK (IKKα), FOS, IKBKB (IKKβ), IRAK1, IRF3, IRF5, I RF7, JUN, MAP2K1 (MEK1), MAP2K3 (MEK3), MAP3K7 (TAK1), MAPK1 (ERK2), MAPK14 (p38ALPHA), MAPK3 (E RK1), MAPK8 (JNK1), MYD88, NFKB1, NFKBIA (IκBα, AD3), RELA, RIPK1, SPP1, TBK1, TICAM1 (TRIF), TNF, TRAF3, TRAF6,CCL3 (MIP-1A), CCL5 (RANTES),CD40 (TNFRSF5), CD80, CD86, CXCL10 (INP10), CXCL11 (I-TAC,IP-9), CXCL9 (MIG), IFNA1, IFNA2, IFNB1, IL12A, IL12B, IL15, IL1B, IL6. AIM2.
NOD-Like Receptors & Signalling Molecules	CARD9, CASP1 (ICE), HSP90AA1, MEFV, NLRP3, NOD2 , OAS2, PSTPIP1, PYCARD (TMS1, SC), PYDC1 (POP1), SUGT1. IL1B, IL18, CYLD
Signalling Downstream of RIG-I-Like Receptors	ATG5, AZI2, CASP10 (MCH4), CASP8 (FLICE), CHUK (I KKα), DDX3X, FADD, IKBKB (IKKβ), IRF3, IRF7, MAP3K 1 (MEKK1), MAP3K7 (TAK1), MAPK14 (p38ALPHA), MAP K8 (JNK1), MAVS, NFKB1, NFKBIA (IκBα, MAD3), PIN1, RELA, RIPK1, TBK1, TNF, TRADD, TRAF3, TRAF6, CXCL10 (INP10), IFNA1, IFNA2, IFNB1 IL12A, IL12B,CXCL8.DAK, DDX58 (RIG-I) HX58 (LGP2), IFIH1 (MDA5), ISG15 (G1P2), TRIM25.
Type I Interferon Signalling	IFNA1, IFNA2, IFNAR1, IFNB1, STAT1 APOBEC3G, IL15, ISG15 (G1P2), MX1, TLR3.

Table 2.2: Shown above is the full list of gene targets tested using gene specific qRT-PCR from Qiagen. The list is available from Qiagen Pathway Central accessed 05/05/2017, and shows functional categories alongside lists of gene targets. The primer sequences for these gene specific assays are proprietary, and designed for use with SYBR green real-time PCR assay.

2.4 Next-Generation Sequencing (NGS)

2.4.1 RNA Integrity Analysis with Bio-Analyser

RNA integrity scores and concentration were assessed using the Agilent 2100 Bio-analyser (Agilent Bioscience). The RNA Pico kit assay chips (Agilent Bioscience) were used to assess picomolar concentrations of RNA and assess the integrity by analysing relative levels of 16S and 28S rRNA using microfluidic chipsets. RNA integrity score (RIN score) of more than 7 was required for use in downstream PolyA selection, library preparation and running on the HiSeq 2500.

2.4.2 DNA removal

RNA samples for analyse with a HiSeq 2500 (Illumina) were initially treated using Turbo DNase (Ambion) and cleaned up using the RNeasy MinElute Kit (Qiagen) as described in section 2.3.4. Following DNase treatment and clean-up the purified DNA free RNA was measured for integrity and concentration using an Agilent Bio-analyser (Agilent Bioscience) as in section 2.4.1.

2.4.3 PolyA Selection

For RNASeq analysis, PolyA selection was performed on DNA free RNA samples to enrich the sample for poly-adenylated RNA species from both host and virus. PolyA selection was performed using the NEBNext PolyA Magnetic Isolation kit (New England, Biolabs). This system utilises magnetic beads coated with oligo$(dT)_{25}$ able to randomly bind to poly-adenylated RNA. After wash steps carried out in accordance with manufacturer's instructions the PolyA selected RNA samples were eluted in RNase free water and analysed using the Agilent 2100 Bio-analyser (Agilent Bioscience) as shown in section 2.4.1 to confirm the lack of rRNA in the samples. The PolyA selection

generates a sample enriched for mRNA species representative of the abundance in the original sample, with non-poly-adenylated RNA reduced or removed.

2.4.4 Illumina Library Preparation

ScriptSeq version 2.0 (Illumina) protocol was followed in accordance with manufacturer's instructions to generate an Illumina library preparation for the HiSeq 2500 (Illumina) Library preparation hybridises fragments of RNA in samples with DNA adapters corresponding to barcoded indices. This allows mixing of samples and loading into separate lanes of a HiSeq 2500 instrument. The barcoding then allows samples to be deconvoluted by bioinformatics. Reverse transcription followed by 14 rounds of amplification were performed with the ScriptSeq version 2.0 chemistry, and samples purified using Ampure magnetic beads (Beckman Coulter) to both normalise levels of adapter labelled cDNA fragments, and to remove reagents which may interview with downstream processing. DNA fragments were eluted in nuclease free water (Ambion), concentration assessed with the Qubit (Thermo Scientific) using the DNA High Sensitivity Kit (Thermo Scientific). Fragmentation was analysed using the AATI Fragment Analyser (Advanced Analytical). Fragment concentration was normalised and validated by qRT-PCR using the Illumina Library Quantification Kit (Kapa).

2.4.5 HiSeq 2500 and data capture.

In accordance with Illumina sequencing protocols utilised by the Centre for Genomics Research, University of Liverpool, template DNA was denatured and loaded at 12 pico-molar concentration into a HiSeq cartridge in the loading port, and 1% PhiX internal control viral RNA was added as a spiked-in control. 6 samples/indices were used per HiSeq 2500 lane to maximise data capture and read depth. Version 4 chemistry was used generating 2 x 125 base pairs, paired-end reads. While up to 96 samples may be added it was decided

in consultation with specialists working in CGR that 6 samples permitted read depths in excess of one hundred million reads per sample.

2.5 Protein Techniques

2.5.1 Protein Extraction with Laemmli buffer

Laemlli buffer is a denaturing buffer commercially available from Sigma-Aldrich. The buffer contains 0.001% Bromophenol blue, 10% Beta-mercaptoethanol (BME), Tris-HCl, NaCl, non-EDTA protease inhibitors (Sigma-Aldrich) and 2-4% sodium dodecyl sulphate (SDS). Laemlli buffer is used to denature proteins, destroy cellular material, inactivate proteinases and permit direct analysis by western blot and SafeStain (Thermo Scientific) staining on gels. The complication with this buffer system is the difficulty assessing exact protein concentration by colorimetric methods due to the presence of bromophenol blue. For use with cellular material infected with ACDP HG4 infectious agents local risk assessment demands that Laemmli buffer is used, and that it is supplemented with an additional 10% SDS (Sigma-Aldrich). For the extraction of protein from cell monolayers, supernatant was removed and cells washed once with sterile PBS. Between 500µl and 1ml of Laemmli buffer was added to the cell layer. The flask or culture plate was then agitated for between 1-2 and minutes to allow time for full cell lysis, and lysates collected into screwcap tubes. Lysates were then heated to 95°C for a minimum of 10 minutes and a maximum of 20 minutes.

This heating step aids in the denaturation of protein and reduces disulphide bonds in the presence of reducing agents such as BME. This also aids in the thermal inactivation, as well as chemical inactivation of infectious agent.

2.5.2 SDS removal by precipitation

Mass Spectrometry is complicated by the presence of ionic detergents such as SDS in samples. Removal of SDS is difficult and the most effective method

is the precipitation of protein and washing of protein pellets to remove excess SDS. The SDS-PAGE Clean-up Kit (GE Healthcare) was employed.

300µl Precipitant per 100µl lysate was mixed and incubated on ice for 15 minutes. A volume of Co-Precipitant equal to the amount of Precipitant was added and incubated for 15 minutes at room temperature. The precipitated protein sample was then centrifuged at ~12,000 x g for 10 minutes at room temperature and supernatant decanted. A small white pellet of protein was visible, which was dispersed in 25µl of proteomics grade water (Waters) by gentle agitation. 1ml of pre-chilled Wash Buffer and 5µl Wash Additive was added, samples were then vortexed aggressively until the protein pellet had dispersed into the wash buffer and then incubated at -20°C for 1 hour.

After centrifugation for 10 minutes at ~12,000 x g the supernatant was decanted and the pellet allowed to air dry for between 5 and 10 minutes. Reconstitution of the pellet was performed with LDS or RIPA buffer.

2.5.3 Protein quantification by RC DC Assay

For protein quantification assays protein samples were precipitated as in 2.5.1 and suspended in RIPA buffer supplemented with EDTA-Free Protease Inhibitor (Thermo Scientific). The RC DC assay (Bio-Rad) was chosen due to the tolerance of the assays measurement for the presence of interfering compounds such as residual SDS. The assay was performed in accordance with manufacturer's protocols and data captured using a spectrophotometer (Molecular Devices) set to read at 750nm wavelength and default settings. Blank assays were performed as for protein assays in accordance with manufacturer's instructions, however consisted only of RIPA buffer.

2.5.4 Western Blot

Western blot analysis was employed to identify bands corresponding to the size of proteins of interest in cell culture lysates. Chemiluminescent western

blot analysis was selected as the most appropriate analysis due to the availability and equipment.

2.5.4.1 Protein Gels

Proteins lysates were generated using the method described in 2.5.1. Lysates were centrifuged at maximum speed or 20,000 x gg at room temperature for 10-15 minutes to sediment insoluble proteins. Precast NuPAGE gels (Thermo Scientific) were unpackaged and rinsed with deionised water to remove the preservative fluid containing 0.0001% sodium azide. Precast gels were loaded into the X2 vertical gel tanks (Thermo Scientific) and 1x running buffer was made by 20 fold dilution of 20x MOPS Buffer concentration (Thermo Scientific). 500µl Anti-oxidant (Thermo Scientific) was added to the central well of the X2 vertical gel tank, which prevents re-oxidation of proteins during gel migration. 1x running buffer was added to the X2 vertical gel tank a 1000µl pipette used to "rinse" wells of the precast gel prior to addition of sample. This step was important as the liquid density in the wells after rinsing with water causes a differential with the running buffer, and samples often do not settle to the bottom of wells unless this was done. Gel tanks were then sealed and connected to a high voltage power back (Bio-Rad). Run conditions for gels were 200V for 50 minutes. For colorimetric staining using SafeStain (Sigma-Aldrich), gels were rinsed 2 times for 5 minutes with deionised water, submersed in staining solution for 60 minutes, then de-stained overnight.

2.5.4.2 Western Blotting

Gels destined for western blot analysis were rinsed with deionised water for 5 minutes after removal of the gel from the X2 vertical gel tank following the procedure described in section 2.5.4.1. To prepare the gel for blotting, the Blot Transfer Kit (Thermo Scientific) was used. This kit comprises pre-stacked

layers onto which a gel was placed. Filter paper was pre-soaked in deionised water and placed atop the gel. The "top" component of the Blot Transfer Kit was then added and the sandwich compressed using a small rolling pin to remove air pockets in the stack. The completed blotting stack was then added to the IBlot Automated Dry Transfer Machine (Thermo Scientific), and the transfer completed in 7 minutes at variable voltage in accordance with manufacturer's instructions. After the completion of the IBlot programme the transfer stack was carefully disassembled and the nitrocellulose membrane transferred to deionised water. The nitrocellulose membrane was checked for successful protein transfer using a reversible stain called Ponceau Red (Sigma Aldrich), or by looking for the transfer of a visible ladder called HiMark (Thermo Scientific).

The membrane was then incubated for 60 minutes in Blotto (Thermo Scientific) blocking buffer, containing non-fat dry milk and TBS with agitation. After initial blocking the membrane was rinsed once with TBS and then primary antibody against the target proteins of interest was added diluted in Blotto, then incubated for 120 minutes at room temperature or overnight (~16h) at 4°C with agitation. Following this incubation period, the membrane was washed for 5 minutes a total of 3 times with TBS containing 0.05% Tween20 (TBST; Thermo Scientific). Secondary antibody conjugated to horseradish peroxidase (HRP) was diluted to the appropriate dilution in Blotto and then added to the membrane.

The membrane was then incubated for 60 minutes at room temperature with agitation. After this incubation period the membrane was washed for 5 minutes a total of 3 times with TBST. The final wash solution was removed from the blot and the excess moisture removed by shaking and placing the blot

in a fresh container. The blot was then exposed to ECL Prime Detection Reagent (GE Healthcare) for 5 minutes.

Excess detection reagent was removed and the blot membrane placed on a trans-illuminator in the GeneSYS Imaging System (SynGene). Image capture was performed using the GeneSYS Imaging System using the on-board automatic capture protocols.

The software was then used to analyse the resultant blot image to gain quantitative information if available. For quantitation information to be available, a secondary loading control needed to have detected at the primary antibody stage.

2.5.4.3 Western blot antibodies

All primary and conjugated secondary antibodies used in western blots in this project are show in Table 2.3. All antibodies were procured from AbCam, Cambridge, U.K. For western blot analysis all antibodies were diluted in Blotto (Thermo Scientific) at the dilution shown in Table 2.3.

Target	Clonality	Origin Species	Target Species	Test Dilution
IL6	Polyclonal	Rabbit	Human	1:1000
TSG6	Polyclonal	Rabbit	Human	1:1000
SQSTM1	Monoclonal (RabMAB)	Rabbit	Human	1:1000
Phospho-p65	Monoclonal (RabMAB)	Rabbit	Human	1:250
Ikkβ	Monoclonal (RabMAB)	Rabbit	Human	1:250
GAPDH	Polyclonal	Rabbit	Human	1:1000
α-tubulin	Monoclonal (RabMAB)	Rabbit	Human	1:1000
MMP9	Polyclonal	Rabbit	Human	1:250
IL1β	Monoclonal (RabMAB)	Rabbit	Human	1:250
IRF3	Polyclonal	Rabbit	Human	1:1000
STAT1	Monoclonal (RabMAB)	Rabbit	Human	1:1000
VDAC1	Monoclonal (RabMAB)	Rabbit	Human	1:1000
VIM	Polyclonal	Rabbit	Human	1:1000
Anti-Rabbit-HRP	Polyclonal (Conjugated)	Goat	Rabbit	1:1000
RESTV GP	Polyclonal	Rabbit	Reston virus	1:1000
EBOV VP40	Polyclonal	Rabbit	Ebola virus	1:1000
EBOV GP	Polyclonal	Rabbit	Ebola virus	1:1000

Table 2.3: The table above displays the commercially available antibodies utilised to detect proteins of interest in assays described in section 2.5.4 of the Methods section of this thesis. The list includes the name of the targeted protein in protein nomenclature. The host from which the antibody was derived is shown, where Rabbit indicated Oryctolagus cuniculus, Human indicated Homo sapiens, Goat implies Capra aegagrus. Virus taxonomic names are given as those most recently proclaimed by the international committee on virus taxonomy. Test dilution is also displayed, indicating the appropriate dilution of antibody in a solution indicated in section 2.5.4. Whether an antibody was monoclonal or polyclonal is also shown.

2.5.5 Filter-Aided Sample Prep for peptide preparation

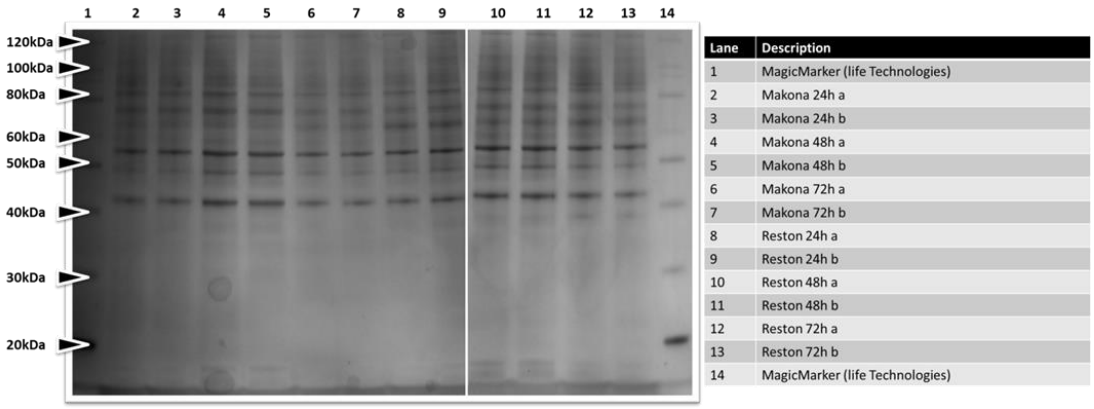
Filter-Aided Sample Preparation (FASP) was a published method which uses microcentrifuge tubes fitted with size-restriction filter columns to remove interfering compounds from protein containing solutions prior to analysis. FASP involves the repeated centrifugation of protein samples in solution by replacing the interfering solution with ammonium bicarbonate (Ambic; Thermo Scientific). A commercial kit was utilised when testing this method (Protein Discovery, UK). Protein lysates containing Laemmli buffer supplemented with an additional 10% w/v SDS (12% total w/v) were diluted 1:5 in Urea-Tris-HCl and spun at 20,000 x g in a microcentrifuge for 15 minutes, or until <50µl remains above the filter, to ensure SDS did not precipitate the centrifuge was kept above 25°C. The filtered centrifuged material was then saved for western blot analysis, to ensure no protein was being lost without realisation. This Ambic wash step was repeated 3 times under the same conditions. After this, 1:10 dilution of iodoacetamide (Protein Discovery) was made in Urea-Tris-HCl buffer and 100µl added to the filter columns. The samples were then agitated on a vortex mixer for 60 seconds and incubated for 20-30 minutes in the dark at room temperature (iodoacetamide is light sensitive). The addition of iodoacetamide alkylates reduced proteins, preventing them from reforming disulphide bonds. The solution was then centrifuged for 20 minutes at 20,000 x g at room temperature. The samples were washed twice with 200µl of Urea-Tris-HCl, with centrifugation steps of 10 minutes at 20,000 x g. The flow through was retained for western blot in all cases. 100µl 50mM Ambic was added to the filter columns and samples agitated for 10 seconds on a vortex mixer, followed by centrifugation for 10 minutes at 20,000 x g. This step was then repeated once more. 75µl Digestion Buffer containing 50mM Ambic and 1:100 dilution of Sequencing Grade

Trypsin (Promega) or Lys C (sigma) was added to filters. Tubes were capped and wrapped in Parafilm (VWR) to prevent evaporation, and incubated for 18h at 37°C. After this incubation the filter columns now contain digested peptides, thus will pass through the filter on centrifugation. 50µl 50mM Ambic was added to the filters and the filters were centrifuged at 20,000 x g for 10 minutes. The filtrate was retained in the tube and the Ambic wash step was repeated again. 50µl 0.5M Sodium Chloride solution (Sigma Aldrich) was added to the filter column and the centrifugation performed at 20,000 x g. The final volume of peptide was thus ~225µl, containing ammonium bicarbonate, sodium chloride and peptides. 2.5µl 1M TFA was added to the filtrate to acidify the peptides and the solutions removed by evaporation using a SpeedVac. Solutions were dried and suspended in 20µl 0.5% TFA (trifluoroacetic acid) and 5% acetonitrile in proteomics grade water (Thermo Scientific), in preparation for desalting using C18 columns.

2.5.6 In-Gel digestion method for peptide preparation

In-gel digestion was performed on gels prepared as described in section 2.5.4.1. Gels were stained blue using SafeStain (Sigma-Aldrich) and thoroughly de-stained prior to use, allowing visualisation of distinguishable bands. The gel was then placed on a light box to aid visualisation, an example of gels prepared in this way is shown in Figure 2.4. Gels were sliced into 18-20 separate slices using a plastic sterile gel cutter (Sigma-Aldrich) under a MSC Class II, to prevent gross contamination of samples. Slices were then cut into 1mm thick cubes and placed into proteomics grade water (Thermo Scientific) then stored at 2-8°C prior to further processing. Gel fragments were fully de-stained using 1:1 vol/vol ratio of acetonitrile and 50mM ammonium bicarbonate and incubation for 30 minutes with agitation. Gel fragments were then dehydrated in 500µl neat acetonitrile (Thermo Scientific) until gel pieces

turn opaque, white and shrink substantially (~30 minutes). Gel pieces were then spun at 16,000 x g for 5 minutes, and acetonitrile removed with a pipette then air dried for 10 minutes to remove remaining acetonitrile. Fragments were submersed in 50-100 iodoacetamide solution, and fragments were allowed to absorb the solution and then incubated for 20 minutes in the dark. Gel fragments were fully dehydrated again using 500µl acetonitrile, and centrifuged at 16,000 x g for 5 minutes once dehydrated, white and opaque. Gel fragments were then saturated with 1:100 dilution of sequencing grade trypsin/LysC.



Lane	Description
1	MagicMarker (life Technologies)
2	Makona 24h a
3	Makona 24h b
4	Makona 48h a
5	Makona 48h b
6	Makona 72h a
7	Makona 72h b
8	Reston 24h a
9	Reston 24h b
10	Reston 48h a
11	Reston 48h b
12	Reston 72h a
13	Reston 72h b
14	MagicMarker (life Technologies)

Figure 2.4: This figure is an example of a Coomassie stained gel containing samples of Ebola virus infected and Reston virus infected THP-1 derived protein, used for both initial protein quality control and to permit in-gel digestion. MagicMarker (Thermo Scientific) was used as a ladder showing marker sizes of 20kDa to 120kDa, and provides a fluorescence signalling when performing western blot. Each lane contains a different sample, and columns were sliced into 18 different gel slices, and cut into 1mm thick cubes using a plastic gel cutter.

After 30 minutes the gel fragments were checked to ensure trypsin digestion solution was adequately absorbed and after 90 minutes 30-50µl ammonium bicarbonate solution was added to ensure enzyme reactions did not dry out during incubation. Tubes were wrapped n Parafilm (VWR) to avoid evaporation, and reactions incubated for 18h at 37°C with agitation. After this incubation period the peptides were extracted by addition of 100µl 1:2 vol/vol dilution of 5% TFA and neat acetonitrile. This dehydrated the gel fragments and thus removed the peptides into the supernatant. The supernatant was collected by pipette and dried in a SpeedVac (Thermo Scientific). The peptides were then suspended in 0.5% TFA in 5% acetonitrile and proteomics grade water (Thermo Scientific), in preparation for desalting using C18 columns.

2.5.7 Desalting with C18 columns

To remove excess salt present following trypsin digestion reactions C18 reverse-phase stage columns were employed (Thermo Scientific). These columns use activated microspheres to bind peptides and allow them to be washed, and interfering compounds removed prior to further analysis. Each reaction can only bind a maximum of 30µg of protein, and thus initial protein concentrations were used which exceeded this, to allow normalisation of concentration between samples. C18 columns were activated by addition of 200µl 50% acetonitrile in proteomics grade water (Thermo Scientific).

Samples then underwent centrifugation at 1500 x g for 1 minute. This step was repeated and then columns were equilibrated with 5% acetonitrile and 0.5% TFA in proteomics grade water, followed by centrifugation at 1500 x g for 1 minutes; this step was repeated again and the flow through discarded. The columns were then placed in fresh collection tubes and a maximum of 150µl of sample (in 0.5% TFA and 5% acetonitrile) was loaded onto the column and samples centrifuged through the column at 1500 x g for 1 minutes.

The flow through was then recovered and this step repeated again. The bound peptides were then washed with a solution of 200µl 0.5% TFA and 5% acetonitrile 2-3 times with centrifugation at 1500 x g for 1 minute each time, and the flow through being discarded. Desalted peptide samples were finally eluted by addition of 20µl 70% acetonitrile, in a fresh collection tube, and centrifuged at 1500 x g for 1 minute. The flow through was retained in the bottom of the collection tube and the elution process repeated with an additional 20µl 70% acetonitrile to maximise recovery. The samples were then vacuum tried in a SpeedVac (Thermo Scientific) and suspended in 0.1% TFA and 3% acetonitrile ready for further processing.

2.5.8 Final method used for Containment Level 4 peptide preparation

After comparing methods of protein lysate clean up and peptide digestion the protocols shown in this section (2.5.8), sub-section 2.5.4.1 and sections 2.5.6 and 2.5.7. For protein extraction the RNA/DNA/Protein AllPrep Mini Kit (Qiagen) was used, Cell lysates in buffer RLT were collected and 600µl 70% ethanol added to precipitate nucleic acid. Lysates were spun through an AllPrep RNEasy Column (Qiagen) at 8000 x g for 1 minute. The reserved liquid in the collection tube after spinning was then transferred to a fresh AllPrep collection tube (Qiagen) and 1.2ml of AllPrep Protein Precipitation Buffer (Qiagen) was added and the samples mixed vigorously for 60 seconds. The precipitation reaction was performed at room temperature or 30 minutes. After precipitation a white precipitant formed and the sample was centrifuged at 20,000 x g in a microcentrifuge tube for 10 minutes at room temperature. The supernatant was carefully decanted without disrupting the protein pellet and the pellet washed with 1ml of 70% ethanol. The washed sample was then centrifuged for 5 minutes at 20,000 x g at room temperature. Again the supernatant was carefully decanted, and the samples left to air dry for a

maximum of 10 minutes to remove residual ethanol. 8mg DTT (Qiagen) was added per 1ml of dissolution buffer ALO (Qiagen). 50µl of ALO containing DTT was then added to the dried protein pellet, and the pellet dissolved in the buffer with vigorous mixing for 60 seconds. The sample was heated to 95°C for 5-10 minutes to allow full denaturation of the protein and encourage it to dissolve in buffer ALO. Buffer ALO is similar to buffer LDS used in other SDS-PAGE protocols. 30µ of protein in buffer ALO was loaded onto a 2mm 10 well NuPage gel with a well capacity of 60µl, in order to maximise protein loading level, and the sample analysed as discussed in sub-section 2.5.4.1. The processed sample was used to measure protein concentration using the BCA Assay (Thermo Scientific). Protein concentration was calculated and a 1:1 ratio mass/mass of protein lysate from "Light" and "Heavy" cell cultures was mixed and processed via the SDS Removal Kit (GE Healthcare) by precipitation. Proteins were reconstituted over 4 hours at room temperature in LDS buffer, and loaded onto a SDS-PAGE gel. Bands were excised from the gel and in-gel digestion performed. Digested peptides were further cleaned up using C18 reverse phase columns (Thermo Scientific), and vacuum dried. Peptides were resuspended in buffers compatible with mass spectrometry and the mass spectrometry performed by Dr. Stuart Armstrong, University of Liverpool, using methods described in sections 2.5.9 and 2.5.10.

2.5.9 Nano-flow liquid chromatography (NanoLC) and mass spectrometry (MS/MS)

A nanoflow liquid chromatographer (NanoLC) was used to analyse peptides prior to mass spectrometry. The Ultimate 3000 nano system (Dionex/Thermo Scientific) was coupled to a Q-Exactive mass spectrometer system (Thermo Scientific). A Nano-Trap column (Acclaim PepMap 100) was used to elute peptides onto an analytical column (Easy-Spray PepMap RSLC) filled with 2µm C^{18} 100 Å micro particles, to allow peptide binding. This analytical column was fused to a nano-electrospray emitter (Dionex). The system was operated at a stable temperature of 35°C. The buffer system for nano-flow chromatography consisted of 0.1% v/v formic acid (Buffer A) and 80% acetonitrile in 0.1% formic acid (Buffer B). The peptides were separated on a linear gradient of 3.8% to 50% Buffer B over 90 minutes at a nano flow rate of 300nl/minute. The Q-Exactive instrument was operated in a data-dependent mode; this means the instrument uses data from a survey scan of the peptide flow to select peptides which meet predetermined rules for fragmenting and further analysis in MS/MS analysis. Survey scans were acquired at a resolution of 70,000 and MS/MS scans at a resolution of 17,500. Up to 10 top abundant isotope patterns with charge states of +2, +3 or +4 from the survey scan were selected for fragmentation by high energy collision dissociation, with normalised collision energy of 30. The time setting for survey scans was set to 250ms and MS/MS scans were set to 50ms. To prevent the repetitive sequencing of peptides, those with sequencing information were dynamically excluded by the instrument software. A diagrammatical representation of this process is shown in Figure 2.5. A schematic of the analysis process and the principles of peptide sequencing and identification are shown in Figure 2.6.

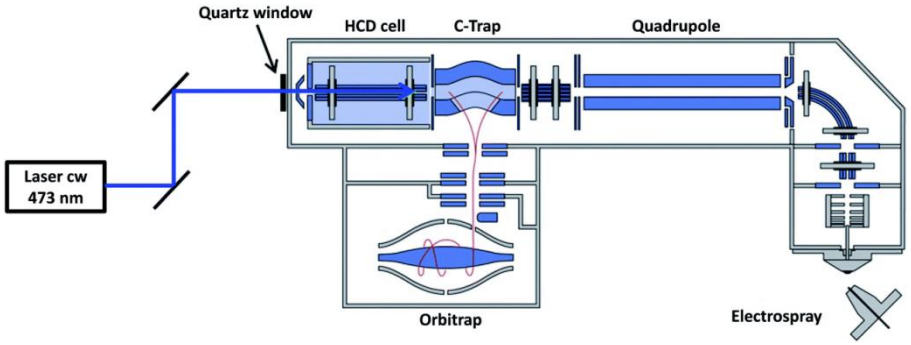


Figure 2.5: Diagram of a Q-Exactive mass spectrometer. This diagram was published by Girod et al. in the Analyst, a Royal Society of Chemistry journal (183). The diagram depicts electrospray ionisation, where peptides are dispersed on a elution gradient by nanoscale liquid chromatography, and sprayed as a fine ionised mist over the intake aperture of the instrument. The instrument pulls single charged peptides through a quadrupole. The quadrupole consists of four charged magnetic rods forming a cylinder. When a peptide of sufficient charge enters the quadrupole it is streamed into the rest of the mass spectrometer, if of high, low or no charge, the peptide trajectory is uncontrolled and analysis of this peptide is not continued. Charged peptides enter the C-trap, which provides a direct current voltage, trapping the peptides in the chamber and forming packets of peptides for injection into the Orbitrap. The Orbitrap causes peptides to oscillate around electrodes in a manner dependent on mass and charge. In this way the peptides are separated by mass-to-charge ratio. The instrument then performs a survey scan and peptides meeting predefined characteristics are injected into second chamber for high energy collision discussion by colliding peptides with gases, and further fragmenting them. In this way peptides with differential mass to charge ratios in the first round of MS can be analysed again in a process called tandem MS, or MS/MS. Fragmented peptides are then analysed and sequenced. In SILAC proteomics, the mass-to-charge ratio is affected by different isotopic labelling in vitro during cell culture and this further mass spectrometric analysis allows a relative comparison between different samples combined and analysed simultaneously.

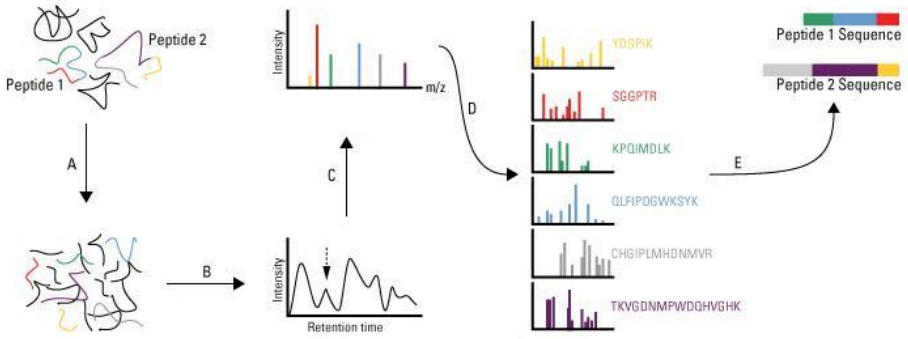


Figure 2.6: A schematic diagram of proteomic analysis adapted from data from Thermo Scientific available from www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/quantitative-proteomics.html Accessed: 02/05/2017. (A) Peptides within proteins are shown, which are separated by fragmentation into separate isolated peptides by digestion, (B) further separated by retention time by nanoscale liquid chromatography, then analysed on a mass spectrometer. (C) The mass spectrometer identifies peptides by their mass to charge ratio, profiling the identities of their constituent amino acids and comparing them to a database, quantifying them using the intensity of their detection signal. (D) Most abundant peptides meeting predefined criteria are then further fragmented by high energy collision dissociation, and peptide characteristics are matched against a database of known peptide associations by bioinformatics analysis.

2.5.10 SILAC-Proteomics

Stable isotope labelling with amino acids in cell culture (SILAC) is a method of metabolically altering the mass-to-charge ratio of proteins and peptides produced by cells *in vitro*, but introduction of a "heavy" variant amino acid in place of its natural form. Over several cell doublings, and in the absence of the natural form of amino acid, the "heavy" isotype amino acid is incorporated into any proteins translated in the cell. This process allows cell cultures containing "heavy" isotype proteins and with normal or "light" isotype proteins to be mixed and analysed by mass spectrometry simultaneously. In this way all processing bias is removed as both samples were processed in synchrony. For SILAC proteomics analysis RPMI media without amino acids was acquired (Thermo Scientific). To make "Heavy" isotype media, 50mg 13C6 L-Lysine-2HCl (Thermo Scientific) and 50mg L-Arginine-HCl (Thermo Scientific) was added to the media and the media filtered through a 2µm sterile filter. Complete "Heavy" media was then made by removal of 50ml media and replacement with 50ml heat inactivated FCS (Thermo Scientific). For "Light" or normal media, 50mg L-Lysine-2HCl and 50mg L-Arginine-HCl was added, and 50ml FCS replaced media to form complete RPMI for the analysis. THP-1 cells were cultured in suspension in either "Heavy" or "Light" media for at least 8 cell doublings or 4 passages. After this period cells were differentiated as described in section 2.1.3.

Differentiated cells were then infected at an MOI of 5 and lysed in Laemmli buffer at 24h, 48h and 72h. Lysates were sampled for BCA analysis and mixed at a 1:1 ratio before being processed as described in section 2.5.8. A schematic representation of this process is shown in Figure 2.7.

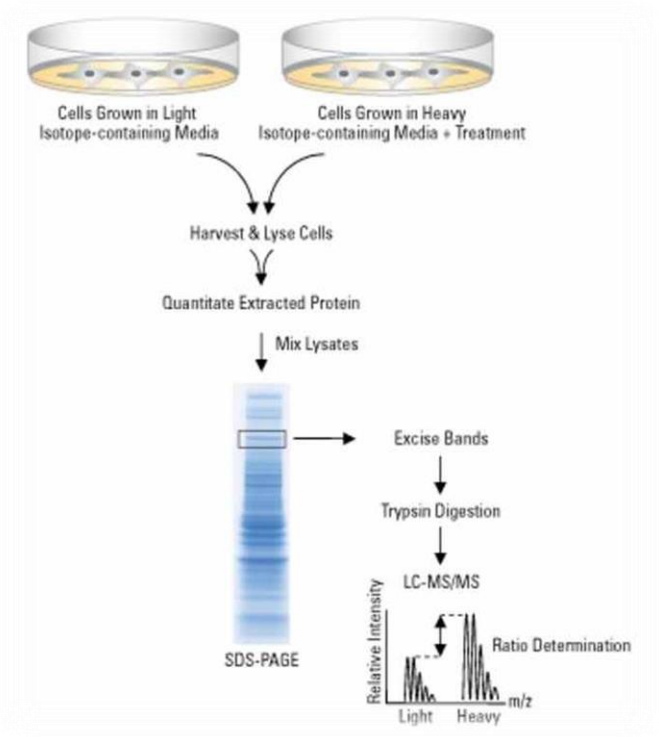


Figure 2.7: A schematic diagram of SILAC Proteomics adapted from data from Thermo Scientific available from www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/quantitative-proteomics.html Accessed: 02/05/2017. Cells were grown in either "Light" media containing normal isotype amino acids or "Heavy" media containing isotopically altered amino acids containing additional carbon atoms, therefore adjusting their mass-to-charge ratio. After harvest and lysis, samples were quantified and mixed at a 1:1 mass/mass ratio prior to further processing. In this way, all processing bias is removed, as samples for comparison were processed in synchrony. Mixed lysates were analysed by SDS-PAGE and bands excised to remove gel slices. Slices were cut into 1mm cubes and in-gel digestion is performed. Purified and cleaved peptides were then analysed on a nanoscale liquid chromatographer and peptide mass to charge ratio analysed by Mass Spectrometry. Peptides were further analysed by a second round of mass spectrometry by collision dissociation of peptides, allowing the identification of different isotypes of the same peptide, and permitting the operator to assess the relative abundance of "Heavy" and "Light" isotypes of the identified peptides.

2.6 Designing a Model of Filovirus Lifecycle

2.6.1 Plasmid design

All plasmids for the mini-replicons of EBOV and RESTV were generated using sequence data available from NCBI. The Ebola virus Makona strain under designation Ebola virus/H. sapiens-wt/GIN/2014/Makona-Gueckedou-C05 was used to generate the gene sequences for each EBOV protein on separate plasmids. RESTV sequences encoding proteins from the Pennsylvania variant M. fascicularis-wt/USA/1989/Pennsylvania were used to generate gene sequences encoding proteins for each RESTV proteins on separate plasmids. Separate plasmids were used for each gene in the system. Gene sequences were codon optimised for mammalian cell expression, artificially created and cloned into the pMX vector (GeneArt, Thermo Scientific). The sequences of the clone inserts were verified and sub-cloned into pUC57_A388 vectors (GeneArt, Thermo Scientific) using the EcoRI and EcoRV restriction sites. Plasmid sequence verification was performed by GeneArt (Thermo Scientific) and glycerol stocks of transformed bacteria were generated using Top10 chemically competent *Escherichia coli* (*E. coli*, Thermo Scientific). Plasmid maps and insert sequences for plasmids encoding the VP24, VP30, VP35, VP40, GP, NP, and L genes for EBOV or RESTV are available in sub-section S1 of the appendices to this thesis.

2.6.2 *E.coli* transformation and amplification

E. coli were used for bacterial transformation and plasmid amplification. Top10 chemically competent cells were used (Thermo Scientific). Cells were thawed from -80°C storage slowly on ice for 30 minutes, then 2-5µg of plasmid DNA was added to cells. Cells were incubated for a further 5 minutes on ice. Top10 cells with plasmid DNA were then heat shocked at 42° for 60 seconds in a water bath, then returned to ice for another 5 minutes.

200µl sterile Super Optical Media (SOC) was added to the heat shocked cells and incubated with aeration for 60 minutes at 37°C on an orbital shaking platform at 400rpm. Thereafter transformed cells were cultured on LB Agar (Sigma Aldrich) containing 50µg/ml Kanamycin (Sigma Aldrich) for 16-21 hours at 37°C without shaking. If culturing from glycerol stocks prepared by GeneArt (Thermo Scientific), glycerol stocks were added to SOC media for 30 minutes at 37°C with shaking and aeration, whereupon the stocks were added to LB Agar (Sigma Aldrich) containing 50µg/ml Kanamycin (Sigma Aldrich) and incubated as for freshly transformed *E. coli*. Individual colonies were selected and cultured for 21-24 hours in 10ml of LB Broth containing 50µg/ml Kanamycin solution with shaking at 400rpm, and incubation at 37°C.

Further amplification was by addition of 250µl or 1000µl of the overnight 10ml LB broth culture being added to freshly prepared LB Broth containing 50µg/ml Kanamycin, at volumes of 500ml or 2000ml depending on requirements for plasmid concentration. Following 21-24 hours culture at 37°C with shaking at 300rpm, LB broth was transferred to sterile 50ml tubes (Corning) and centrifuged at 5000 x g at 4°C. Broth was removed and cell pellets were either processed as described in sub-section 2.8.3 or frozen at -80°C until required.

2.6.3 Plasmid purification

The HiSpeed Plasmid Purification Kit (Qiagen) was employed to extract and purify plasmid from *E.coli* cells in accordance with manufacturer's instructions. The "HiSpeed" kits replace the essential centrifugation steps at the terminal phase of the extraction process with a silicon filter based method. Pellets of transformed cells containing the plasmid of interest were lysed using a series of buffers. First the lysis buffer containing RNase A was added to the cells, followed by a further proprietary lysis buffer containing SDS.

Finally a neutralisation buffer was added which fully precipitates all cellular and protein material from the buffer. The precipitate was removed by filtration through a fibreglass filter column, and the clarified supernatant was then supplemented with alcohol containing buffers to precipitate nucleic acid.

The supernatant was then bound to a HiSpeed Filter Cartridge (Qiagen) and washed repeatedly with an alcohol based wash reagent. 60ml in total was used to remove contaminants. 10.5ml of Elution Buffer (Qiagen) was added to the filter column and 5ml of ice cold isopropanol was then added. After mixing the solution containing precipitated plasmid was filtered through a HiSpeed Ultrapure Filter Cartridge (Qiagen), washed once with 5ml of Isopropanol, then once more with 2.5ml of 70% ethanol, made using Endotoxin free water (Qiagen). Finally 1ml of endotoxin free water (Qiagen) was added to the column to extract the plasmid from the filter cartridge. The eluent was then returned to the filter cartridge and the elution step repeated to maximise plasmid return. Following this process the plasmid samples were analysed by Nanodrop 2000 (Thermo Scientific) and abundance of DNA calculated from absorbance at 280nm and 260nm. Plasmid samples were stored at -20°C until required for transfections or analysis.

2.6.4 Plasmid sequencing by NGS

Transformation of plasmid into competent cells or amplification of plasmid in cells may result in mutations. To assess plasmid sequences for the presence of any unexpected changes or mutations, samples representing each plasmid encoding the VP24, VP30, VP35, VP40, L, NP and GP for EBOV or RESTV were pooled together and analysed by Illumina next-generation sequencing. 10µg of each plasmid was pooled at a 1:1 ratio with all other plasmids and concentrated to 10µg/µl using ZymoFilter columns (ZymoMax).

Concentrated DNA was added to 0.1ml microcentrifuge tubes (Eppendorf) and sent to the Public Health England, Sequencing Service, Colindale, London, for library preparation and sequencing using a MiSeq NGS System (Illumina). FastQ files were curated using FastQC and reference sequences corresponding to each codon optimised gene sequence were uploaded to GALAXY to act as reference sequences. The BWA tool in GALAXY was used to assemble Illumina sequencing reads using the relevant EBOV or RESTV gene reference as a template for mapping. The resultant BAM file was indexed to create a BAM index file (BAI) and mapping effectiveness was visualised directly using Tablet. This method allowed easy analysis of insert gene sequences, but made it impossible to assess any variation upstream or downstream of the viral gene sequences, located in the plasmid or promoter regions of the constructs. The gene BAM file was analysed using QuasiBAM to look for sequence variation and produce a consensus sequence. The consensus sequence report was produced and compared with the relevant gene reference sequence to look for evidence of variation using MegAlign Pro (Lasergene Version 11.0).

2.6.5 Transfection of mammalian cells *in vitro*

Transfection of cells with plasmids was performed using the TransIT-LT1 transfection reagent (Mirus BIO LLC) in accordance with manufacturer's instructions. TransIT-LT1 was removed from refrigerated storage and warmed to room temperature for 30 minutes prior to reaction set-up. The reagent was vortexed at low speed for 10 seconds to disperse lipid. Plasmids were added to Opti-MEM serum free media (Thermo Scientific) and mixed together, then transfection reagent was added and mixed carefully by repeated pipetting (tituration). To allow complex formation the transfection mix was incubated at room temperature for between 15 and 30 minutes. Transfection mix was then added to cell cultures in a drop wise fashion.

Most transfections were performed in a 24 well format. Requiring 50µl total volume per well, and 1.5µl of TransIT-LT1 per well. After addition of complexes, the culture plate/dish was rocked gently back and forth to disperse the complexes in culture media over cells. Culture plates/dishes were then returned to 37°C and 5% CO2 incubation.

2.6.6 Luciferase assay systems

For luciferase activity measurement the BrightGlo Luciferase Assay Kit (Promega) was used. BrightGlo reagent was reconstituted in 100ml of BrightGlo Assay Buffer (Promega), and mixed thoroughly. Aliquots were then stored at -20°C or were used in assays immediately. The GloMax (Promega) luciferase analyser was primed for 5 minutes to warm up the photo-multiplier tube. BrightGlo reagent was added to wells of a 96 well plate at 1:1 vol/vol with culture media in wells containing transfected cells producing luciferase. BrightGlo reagent was used to analyse luciferase which is contained within cells and not exported. Therefore the reagent contains SDS and DTT in order to lyse cells and release luciferase into the supernatant. The time needed to ensure total cell lysis and for enzymatic activity to plateau is recommended to be 5 minutes at room temperature according to kit instructions. Luciferase assay activity was measured on the primed GloMax reader (Promega) at 0.1 second intervals. Luciferase light units were normalised to control sample results to calculate relative light units. After addition of BrightGlo reagent (Promega) luminescence results were stable for 10 minutes and loose luminescence quickly after that, and can be stably analysed for 2 hours.

2.6.7 MTS cytotoxicity assay

To measure the cytotoxic effects of chemicals and drugs used in stimulation and inhibition assays, the CellTitre 96, Cell Proliferation Assay (Promega) was used. This assay reagent contains a tetrazolium compound [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, and an electron coupling reagent (phenazine ethosulfate; PES). PES combined with MTS forms a stable solution. The MTS reagent was reduced by cells into a formazan compound which was soluble in culture media and was measurable as a strong purple colour change. Unlike MTT assays this assay does not require a solubilisation step, as the coloured formazan product was soluble on generation. A reduction in the intensity of this colour change correlates directly with reduced cell metabolic activity, and was indicative of cell death, or cytotoxicity. The One Step Solution (Promega) was thawed for 90 minutes at room temperature, and 20µl added to each well of a 96 well plate containing 100µl of culture media and the cells to be assayed. The assay plate was then incubated at 37°C with 5% CO2 for 4 hours. After this incubation period the plate was placed on a plate shaker for 5 minutes at 100rpm. The plate was then measured on a plate reader at 490nm.

2.7 Bioinformatics Analysis

2.7.1 NGS data analysis in GALAXY

GALAXY is an online html system providing scientists with a user friendly interface, accessing a library of command line scripted algorithms for the analysis of next-generation sequencing data. All NGS analysis using GALAXY was performed using a secure partition operated by Public Health England which permits the encryption of data, and thus allows the analysis of sensitive datasets, such as those regulated by NHS governance or Schedule 5. Illumina data in a fastq format may be mapped to a reference fasta sequence using BWA. BWA aligns reads from a fastq file to the reference sequence and assembles a mapping index based on this alignment. In this way, BWA is a useful tool for mapping to expected or known targets, but biases the analysis to the intended reference sequence. The BAM output contains mapped reads or unmapped reads. Mapped reads BAM can be indexed to create a BAI file, which rearranges sequence reads in a BAM file so that it can easily be matched to a location on a reference sequence. Indexed BAM files may then be visualised using Tablet, allowing the user a view of the sequencing effectiveness, and whether there were discernible gaps in the sequencing or areas of low read depth which require careful interpretation.

2.7.2 Quasispecies analysis using QuasiBAM.

QuasiBAM is a tool developed by Public Health England to measure population diversity in a BAM file by assembling mapped reads against the reference genome, and looking for nucleotides which appear to vary in reads. Those which appear to be in great abundance up to around 50% of reads may indicate a stable mutation in the virus population.

Smaller percentages of reads showing a variance were deemed minority variants and may be significant if population diversity is of interest. QuasiBAM is used to measure the minor variants and major variants in a population of virus sequenced and also produce a consensus genome sequence, which may be used for further analysis.

2.7.3 Use of the R Environment

R is a free statistical software environment which allows access to a vast library of algorithms and scripted packages to perform statistical and bioinformatic analysis. RStudio version 2.1 was used, and statistical packages for RNASeq were downloaded including limma, EdgeR and DESeq, additional required packages were also downloaded and graphical packages including gplots and ggplots 2 were downloaded to enable to creation of graphs and plots of RNASeq data.

2.7.4 Differential Transcriptomic analysis using EdgeR

EdgeR is a software package in the R environment used to calculate fold change in a transcript comparing one sample to another and modelling read count data to determine significance (140). A .gtf file containing the coding sequence information and a reference genome in a .fasta format is provided to enable mapping. In this case the reference used was the hg19 GRCh38 human genome available from ftp.ensembl.org/pub/release77/fasta/homo_sapiens.

The .gtf file was the transcript annotation for the reference genome and the release 77 of the GRCh38 reference annotation file was used and was available from ftp.ensembl.org/pub/release77/gtf/homo_sapiens. As Illumina HISeq 2500 produces paired end reads, read pairs were counted rather than single reads. First paired end reads were trimmed using CutAdapt (184) to remove primer sequences. Then paired end reads were mapped to the reference genome and annotation data using Bowtie2, a fast read mapper (185).

The resultant mapped read file (SAM/BAM) was then analysed in HTSeq-Count to calculate the number of reads which map to each gene in the annotated genome using the SAM file produced by Bowtie2 and the .gtf file which informs HTSeq-Count where the boundaries of genes are located on the reference genome (.fasta). Finally, the normalised read counts from HTSeq-Count, known as counts per million (CPM) were analysed using EdgeR. As RNASeq data typically does not assume a normal distribution EdgeR uses negative binomial distribution to model read counts and determines whether a difference between two samples was statistically significant. EdgeR also produces a number of quality assessment plots including an assessment of variation/dispersion to determine whether read count biases variance in a particular transcript. Using EdgeR fold change of one sample to another was calculated, and was the p-value of significance was corrected by false-discovery rate. Often in addition to significance a fold change "cut-off" was imposed, which was arbitrarily chosen to be 2. Therefore a fold change of +2 or -2 from the comparator sample was deemed to be a significant change only if the corresponding FDR corrected p-value was <0.05.

2.7.5 Proteomic data analysis using MaxQuant software and Andromeda search engine.

Mass Spectrometry data analysis was performed using the MaxQuant software (version 1.5.5.1)(186). Peptide identification was performed using the built in peptide search engine called Andromeda (187). Analysis was performed using the default settings. The search was set to include variable modifications of methionine oxidation and N-terminal acetylation. The Enzyme specificity for the analysis was set to either LysC or Trypsin depending on the upstream processing which had been employed in protein cleavage. The false discovery rate correction was set to 0.01 for peptide and protein identification.

For SILAC proteomics analysis calculation of relative abundance of peptides was based on "Light" (Lys 0) or "Heavy" (Lys 6) labels. The Andromeda peptide identification search engine was set to include *Homo sapiens*, EBOV and RESTV proteins. Additionally the software included a database of contaminants to determine correct false discovery rates (FDR) and exclude false positive identifications due to protein contamination of other species. Protein groups were considered reproducibly quantified if identified and quantified in both replicates in an experiment.

2.7.6 Gene Ontology analysis

In order to assess the biological significance of gene and protein abundance information, gene ontology analysis can be used to assess the combined change of functional categories. Several tools are available to do this, and each use similar methods to determine the significance of such categorisation.

2.7.6.1 PANTHER

PANTHER is a web based statistical enrichment test. Genes deemed significantly change in analysis by EdgeR were selected and gene identifiers listed in PANTHER, the enrichment test compared the list of genes to the background list (usually the whole human transcriptome) to determine whether there seems to be an over-representation of a particular functional category of gene or protein in the provided list. This comparison with a background list allows significance testing, and a p-value is given to assess the significance of that over-representation.

2.7.6.2 GOrilla

GOrilla is a gene ontology tool which looks for linked categories and functional networks in a list of genes determined to have changed significantly when analysed by EdgeR. The result was that gene lists are annotated with functional categorisation. Default settings were used in this analysis

2.7.6.3 KEGG Pathways

KEGG is a database of functional properties assigned to genes and proteins. KEGG pathways include canonical pathways, and matches genes in a given list to a database of canonical pathways to determine best fit. The canonical pathways with the highest representation in the gene list are then returned.

2.7.6.4 STRING 10.0

STRING version 10.0 allows a list of genes to be assessed for connections, interactions and functional networks. Orphan genes can be removed and sub-networks in a list of significantly changing genes can be identified. Functional categories from the KEGG database as well as gene ontology databases are given, allowing simultaneous analysis of many functional characteristics of the protein. When used the tool was set to the default settings.

2.7.7 Ingenuity Pathway Analysis

Ingenuity Pathway Analysis (IPA; Qiagen) is a commercially available toolkit for the analysis of gene and protein expression data. Fold change data is uploaded and annotated in the software. The software analyses for over-represented canonical pathways, functional groups and disease groups. Additionally, activation score (z-score) was calculated for canonical pathways to determine the level of activation. This pathway z-score was calculating using the cumulative fold increase or fold decrease of the genes aligned to that pathway. In this way the activation of host response elements can be quickly assessed and analysed further. IPA was used to determine canonical pathway activation, functional categories of genes and proteins and was used to assess and predict the activity or activation of upstream regulators and transcription factors.

2.7.8 Single-Site Analysis with oPOSSOM

oPOSSUM is a search engine which uses statistical enrichment to determine whether there the targets of a particular transcription factor or expression regulator are over-represented in a list of genes or proteins. Gene lists were loaded onto oPOSSUM via their web based portal, and compared with 20,000 randomly selected genes from the human transcriptome in order to calculate over-representation. Transcription factors are assigned a fisher score (significance) and activation score (z-score; over-representation).

Those with both a high significance and high z-score are viewed as being predicted to have a role in the transcriptional regulation of a large number of genes/proteins in the provided gene list, and were selected for further analysis and functional investigation.

2.8 Statistical Analysis

Statistical analysis was performed within GraphPad Prism Version 5.0 and Version 6.0, in MiniTab version 16, or in Excel 2010. Several statistical tests used in this thesis are described briefly in this section.

2.8.1 Q-Q Plot

A Q-Q plot (quantile-quantile plot) is a descriptive test to ascertain whether two groups of data share a common pattern of distribution. This also shows whether data follows patterns which fits a normal distribution or not, and thus allows statistical tests to be selected based on parametric or non-parametric methods, or selection of statistical tests which do not assume normality. Two sets of data are plotted against each other, x-y and a 45° line was plotted which acts as a reference for data correlation. In this way a quick estimate of data similarity can also be discerned.

2.8.2 General Linear Models

General Linear Models are employed in multi-variate analysis. Where data contains meta-data (descriptive data; such as temperature, time, virus type, cell type etc…), or many variables which affect the data must be taken into account. It allows multiple models of regression to be tested simultaneously (multi-variate), or as several separate single variable tests (uni-variate), in order to ascertain whether data follows a pattern which was potentially effected by or correlates with each confounding descriptive variable or dataset.

2.8.3 ANOVA

The ANOVA (Analysis of variance) test is conceptually similar to performing multiple two-sample T-tests on two datasets. The ANOVA allows the simultaneous comparison of multiple groups or variables, and tests whether the means of each group or variable are equal or not. ANOVA may also be used for multi-factorial analysis, where several variables may have an effect on data. In this way it applies the same principles as general linear models, and tests whether variables have an effect on the datasets tested.

2.8.4 Spearman's Rank Correlation

The spearman's rank correlation test is a method of identifying correlation between two sets of data, ascertaining whether this is positive, negative and whether this is significant. Unlike the Pearson's correlation test, the Spearman's rank test does not assume data is parametric. As most RNASeq data is non-parametric the Spearman's rank test was chosen to compare datasets across the project to maintain consistency and comparability.

The Spearman's rank correlation is given as an R value, where a result of 1 or -1 indicated where datasets are in perfect correlation with the other, positively or negative, respectively.

2.8.5 Principle Component Analysis

Principle Component Analysis (PCA) involves the transformation of potentially correlating data into linearly uncorrelated values known as principle components by an algebraic orthogonal transformation. The analysis also reduces the number of variables shown, so that entire datasets can be described, rather than individual data points. The first component of a principle component analysis is the most sensitive, and each following component requires further transformation, making the analysis less sensitive. The analysis plots principle components in a 2D or 3D space. The linear distance between points in a principle component analysis plot shows the similarity in the transformed data.

Thus PCA is conceptually an analysis of how similar data is, and is a useful method of modelling datasets to gain a broad view of them, and which variables can be used to separate data.

2.8.6 $TCID_{50}$ Calculation

$TCID_{50}$ (tissue culture infective dose 50%) assays are a method used to calculate the dilution of stock/sample required to cause cytopathic effect in 50% of cells (or culture vessels containing cells in a multi-vessel assay).

A dilution series is performed and the required dilution of virus to kill 50% of cells in a cell culture is found by using a statistical test using the Reed and Muench 50% end-point method (181).

The Reed-Muench method can be applied to any 50% end point experiment, to ascertain dilution required for 50% infective dose, lethal dose or inhibitory dose of a chemical, drug or virus. The reasoning behind using a 50% end point for calculations is that most biological data follows a logistic function, and if plotted on a graph forms a sigmoidal curve with a low and high plateau. The 50% end point therefore gives a more accurate estimate of concentration by avoiding the minimum and maximum response or effect. After calculating the exact dilution required for a 50% tissue culture infective dose, this can be used to control the inoculum of virus used in subsequent experiments through MOI calculations (mode of infection). For example, an MOI of 0.5 $TCID_{50}$ units is equivalent to an inoculum containing half the concentration of virus which was needed to infect 50% of cells in culture. An MOI of 5 $TCID_{50}$ units is equivalent to an inoculum containing 5 times the required inoculum to infect 50% of cells in culture. Similar to plaque assay results, $TCID_{50}$ titres are given as 1x10^(50% dilution) $TCID_{50}$/ml. For example, calculating an end point dilution of $10^{-6.5}$ indicated a virus concentration of 1 x $10^{6.5}$ $TCID_{50}$/ml. All $TCID_{50}$ calculations were performed in accordance with the 1938 paper.

3. Characterising an emerging Ebola virus variant

3.1 Introduction

3.1.1 The original outbreak variant of EBOV

The Yambuku variant was collected during the original outbreak in Yambuku, Zaire (now the DRC) in 1976. Two isolates were successfully retained and investigated. One was called Mayinga, and was isolated from samples collected from a catholic nun working with EVD patients in the outbreak (39). The second isolate was called Ecran (19), from a sister working in the same hospital. Sequencing studies have revealed that EBOV was transmitted from Sister Mayinga to Sister Ecran early in the outbreak resulting in fatal infection. Thus the isolates are closely related temporally, epidemiologically and phylogenetically. The Ecran variant was originally sent to Antwerp, Belgium in 1976, and used in studies thereafter by virologists at Porton Down. The Ecran variant was recently re-investigated to map its distribution and origin (19). The Yambuku-Ecran variant has been used in experiments at Porton Down for over 40 years.

Most of this has involved repeated passaging in Vero cells; Both the Yambuku-Ecran variant and the Makona variant viral inoculums were prepared from supernatants collected from amplification cultures on Vero cells.

3.1.2 Diagnostic findings in the 2014 outbreak of EVD in West Africa

The first reports of EVD in West Africa emerged from the forested eastern regions of Guinea. The initial laboratory response was facilitated and supported by the European Mobile Laboratories (EMLab), working closely with Medicins sans Frontiere (MSF) to deliver emergency care in the region. The first centre for laboratory diagnosis was established in Guékédou, around 70km south of the nearest major settlement known as Kissidougou. The EMLab provided laboratory response in Guékédou from March 2014 until March 2015. Finalised reports showed the laboratory tested 4719 samples collected at admission to the Guékédou ETU from 2741 patients during this period. 79% of patients attended hospital or the Ebola Treatment Centre (ETU) and 21% were classified as community/suspicious deaths. qRT-PCR analysis was used to confirm EBOV infection, and showed 1231 of the suspected EVD cases were positive of those treated in hospital or ETU, and 281 of the community deaths were positive by qRT-PCR. The full dataset is summarised in Figure 3.1, published by the EMLab Ebola response team and listed under associated publications (188).

Characteristic	EVD Suspected Cases in Hospital			Community Deaths		
	Overall	EBOV RT-PCR Positive	EBOV RT-PCR Negative	Overall	EBOV RT-PCR Positive	EBOV RT-PCR Negative
Individuals/Total	2178/2178	1231/2178	947/2178	563	281/563	282/563
Females/Total	1135/2157	645/1228	490/929	260/545	136/271	124/274
Median Age	30 (18–44)	30 (19–45)	30 (18–42)	37(25-55)	35 (23-53)	40 (25–56)
Malaria (positive/total)	541/1937	261/1091	280/846	Not tested	Not tested	Not tested
Fatal outcome (Fatal/total)	769/2049	719/1205	50/844	563	281	282

Figure 3.1: Summary of the European Mobile Lab database showing the number of positive cases, sex ratio, age, outcome and whether the patient had a corresponding positive result from the Malaria rapid test performed in synchrony with Ebola virus qRT-PCR testing. All data was collected live during outbreak reporting on a daily basis from EMLab, Guékédou. Data was collated by Kerber et al [189] and a full analysis of the data is available in the attached publication.

3.1.3 Case fatality rates

Classification of 2178 patients by age, malaria positivity, EVD and fatal outcome revealed case fatality rates increased in patient subsets with simultaneous EVD and malaria presentation in an age dependent manner.

Patients aged 0-4, 15-44 and >44 with detectable plasmodium co-infection did not show increased case-fatality rates in comparison with those with EVD alone. In contrast patients aged between 5 and 14 showed an increase from ~40% CFR in EVD patients without malaria positive result to ~65% CFR in patients with detectable plasmodium co-infection. Additionally this age group displayed the highest levels of malaria positivity in the patient cohort. This data was generated through emergency diagnostic data produced during the acute stages of the outbreak. I was personally involved in this data collection and am recognised on the corresponding paper Kerber et al [189]

3.1.4 Characterising the host response to the Makona variant

Investigations of the biological differences in Makona compared with other isolates of EBOV have found differences in the ability of Makona and Mayinga to enter human-derived cell lines (189), and Makona has demonstrated adaptation to human cell entry (190). Work in the *in vivo Macaca fascicularis* animal model showed a shift in pathogenicity, where disease course to fatality was delayed in Makona infected animals in comparison to the Mayinga archetypal isolate from Central Africa (Yambuku 1976) (191).

The possibility remains that infection with Makona may result in differences in patterns of activation or suppression in infected cells compared with viruses from Central Africa. Profiling cellular changes during viral infection successfully identified differences in cells infected with highly pathogenic H1N1 influenza virus compared with a lower pathogenicity variant (192)

3.1.5 Chapter aims

The first aim of this project was to identify the host response signals specific to the emergent *Ebola virus*, Makona variant compared to other variants of EBOV. The objective was to compare the Makona variant to an archetypal variant originating in Central Africa (1976), by contrasting the genome sequence of the Makona variant, by investigating the host response to the Makona variant virus in a human patient and by directly comparing the cellular response to both the Makona variant and a Central African variant in cell culture to identify host response signatures unique to the Makona variant.

3.2 Chapter methods

3.2.1 Phylogenetic analysis of Makona strain

All samples collected through emergency diagnostic activity were dispatched to the Bernard-Nocht Institut to form a biobank of samples to permit characterisation and study of the EBOV variant causing the outbreak affecting West Africa. Samples were re-extracted in the bio-safety level 4 laboratories of the Bernard-Nocht Institut, Germany, and sent to the University of Liverpool for next generation sequencing (NGS) using an Illumina 2500. 6-10 samples were indexed per Illumina sequencing run to maximise data acquisition and read depth. Many samples acquired were suspected to be of poor quality due to extended periods in undesirable storage conditions, to compensate for this RNA was checked for quality using bio-analyser prior to initiation of sequencing pipeline. The original testing, specimen collection and specimen handling at high containment were performed with my involvement, recognised in the corresponding authorship of Carroll et al 2015.

In total 179 RNA extracts from patient samples were successfully sequenced. Next-generation sequencing was performed at the Centre for Genomics Research as part of the EMLab Consortium collaboration. 179 EMLab sequences collected as described above were downloaded from Genbank alongside sequences from Yambuku (Mayinga and Ecran) collected first in 1976, sequences from Boende and Lomela in 2014, and sequences from outbreaks in Kikwit and Gabon. Sequences were trimmed using MEGA6 and aligned using MAFFT in the GALAXY environment. Phylogenetic analysis was performed using RAxML, performing a maximum likelihood analysis by first creating a neighbour joining tree, and re-analysing by maximum likelihood.

A wide variety of sequences are available from March 2014 to March 2015 which have been collated and analysed in an open source platform known as NextStrain from a number of published sources (22, 193-197). The platform has allowed easily accessible phylogenetic information to be obtained on isolates from the West African outbreak in Sierra Leone, Liberia and Guinea.

3.2.2 Analysis of viral RNA from a case of EVD imported into the UK

RNA sampleswere collected on days 4-12 of illness by the clinical team at the Royal Free Hospital, London, UK. These had the highest viral load and were selected for next-generation sequencing. Sequencing reactions were performed in conjunction with Andrew Bell and Kuiama Lewandowski, Public Health England, UK. The patient samples were sequenced according to local protocols and data captured on an Illumina MiSeq next-generation sequencing platform. Viral sequences were mapped to the accessioned sequences of the Makona-C05 isolate.

Transcript abundance data was generated using Cufflinks in the GALAXY environment and further analysed using Ingenuity Pathway Analysis to identify biologically relevant activity. Transcripts with FPKM values calculated for a majority of time points in the time course analysis were included in further analysis.

All work has been reviewed by the NHS National Ethical Committee, and North West Regional Ethics Committee, approved Research Ethics number 15/NW0744, Full documentation is in appendix I.

3.2.3 Transcriptomic changes in EBOV Makona and Yambuku in A549 cells

Both the EBOV Makona and EBOV Yambuku-Ecran variants were titred by $TCID_{50}$ on Vero cells and A549 cells were infected with an equal MOI (0.5 $TCID_{50}$/cell) in accordance with sub-section 2.1.4 and 2.1.5. A549 cells were selected for this study because of their characteristically strong antiviral response as a useful cell line to model differences between the Makona and Ecran variants and their interactions with the interferon response.

The abundance of viral genome was compared by qRT-PCR in accordance with section 2.3 and analysed using a repeated measure General Linear Model with Greenhouse-Geisser correction.

Duplicate infections were performed in parallel and qRT-PCR assays were performed in triplicate for the purposes of statistical analysis. Transcriptomic analysis of RNASeq data was performed using HTSeq-Count as described in Subsection 2.6.5. This analysis directly counted the trimmed reads mapping to mRNA and normalised this value per million bases of reference genome to give CPM.

To compensate for this large range in total read counts acquired and avoid biasing the analysis because of additional read depth in certain datasets, normalisation was performed to account for size factors using EdgeR. After normalisation CPM was analysed using EdgeR to identify significant differences in transcript abundance. This analysis uses general linear models to identify significant changes and calculates corrective false-discovery rates.

3.3 The Makona variant genome sequence is stable and similar to other variants from Central Africa.

3.3.1 Viral load data correlated with patient outcome

Data from qRT-PCR measurement of viral genome in patient samples collected in Guékédou, Guinea was contrasted with patient fatal/non-fatal outcome, revealing a clear correlation between low qRT-PCR cT value and poor outcome. Patients with a cT of between 14 and 19 had a CFR in excess of 50%. In contrast samples showing a cT of >25 suggested the patients had a CFR of <30%. This was a clear indication that viral load determines patient outcome.

3.3.2 Makona variant isolate sequences cluster together and separate from the Central African isolate sequences

To determine whether the Makona variant showed distinctive genomic changes in comparison with other isolated EBOV variants, genetic analysis was performed by next-generation sequencing.

Phylogenetic analysis comparing 179 Makona isolate sequences, and sequences from viruses isolated in Central Africa shows that the West African EBOV sequences may be separated from Central African sequences at a resolution of 0.006 substitutions/1000 nucleotides (Figure 3.2). This analysis included those sequences from recent isolates derived from Lokolia and Boende in 2014.

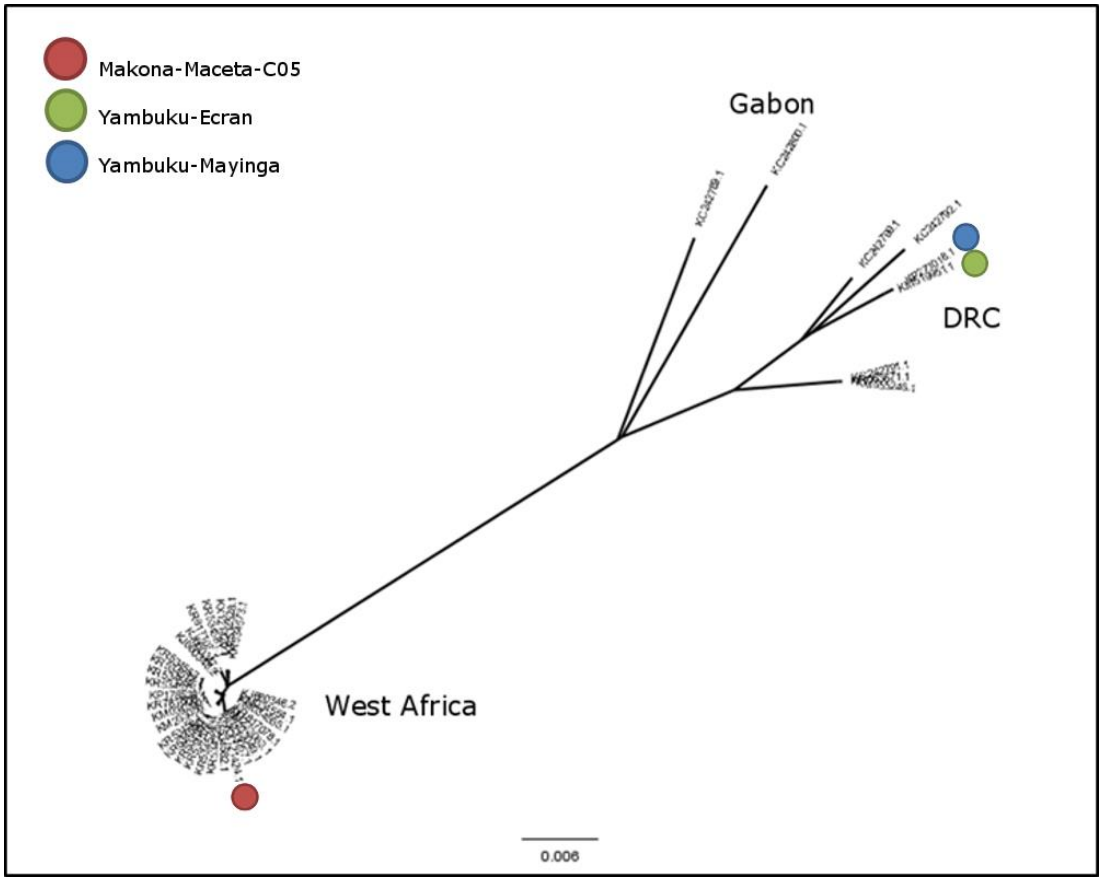


Figure 3.2: Maximum likelihood phylogenetic analysis of 179 genomes collected from the West African outbreak in Guékédou produced using FigTree with alignment and ML analysis performed using RaxML using the General reversible time (GRT) model of phylogenetics. Variant sequences from guinea were compared with sequences of EBOV collected from previous outbreaks in the Democratic Republic of Congo (DRC) and Gabon. The scale shows a nucleotide substitution of 0.006 per thousand nucleotides indicated by the size bar below the tree. An un-rooted tree was selected to show the clustering of West African sequences distinctly from Central African sequences. Sequences of the Makona isolate C05, and sequences of the Yambuku-Ecran and Yambuku-Mayinga isolates from 1976 are annotated in the tree to display similarity.

3.3.3 The genome sequences of the Makona and Ecran variants were similar

To evaluate whether there were distinctive genome changes in the Makona variant or Yambuku variants, sequences of stocks of these viruses were compared with accessioned sequences of other EBOV isolates, stocks were produced in the Porton Down Containment level 4 suite and have been passaged on authenticated Vero cells. Next generation sequencing data from the virus stocks of Ecran, Mayinga, Makona and Kikwit were mapped to the NCBI EBOV RefSeq. Sequences were aligned using MAFFT by maximum likelihood analysis following a neighbour joining alignment method and trimmed for 3' and 5' terminal regions. A phylogenetic tree was resolved at a scale of 0.002 distance/substitutions. Pair-wise distance was computed using default settings in the MEGA6 software. The calculated divergence in nucleotide sequence in Makona compared with Ecran or Mayinga was 3% overall (difference between Ecran and Mayinga was ~0.03%). Interpretation of these combined analyses suggested that EBOV was not highly variable and that Ecran, Mayinga and Makona have similar genome sequences. This is surprising but reassuring; similarity between these isolate sequences indicates that despite the lengthy passage history of Ecran and Mayinga as lab-adapted strains, genome sequences align well with the arguably more clinically relevant sequence of Makona. The similarity in sequence between Makona and central African variants (such as Ecran) suggested that the role sequence mutations may play in determining differences in the host response would be easier to identify.

3.4 Longitudinal analysis of an EVD case presenting in the United Kingdom.

3.4.1 Importation of a case of severe EVD into the UK

The outbreak progressively worsened by the end of 2014 with cases in Sierra Leone especially high (57). The risk of an imported case was suggested to be rising (198), exacerbated by increasing numbers of humanitarian and healthcare workers which were deploying to the affected areas. In December 2014 a healthcare worker returned to the UK with acute febrile illness and general malaise. On presentation to hospital the patient had a developing erythematous rash and was deteriorating (199). qRT-PCR for EBOV genome was performed in conjunction with the Rare & Imported Pathogens Laboratory (Days 1-2) and Virus Reference Department, Colindale (Days 3-28). The patient's initial sample was tested and viral genome copies/ml of blood calculating using the standard curve method. The patient showed 1.43 x 10^7 copies/ml (cT 23.9) viral genome on presentation. On day 2 this rose to 8.19 x 10^7 copies/ml (cT 21.2).

The patient was treated with Brincidofovir and convalescent plasma from EVD survivors in West Africa. By day 5 symptoms worsened, the patient developed severe watery diarrhoea and respiratory failure, did not develop haemorrhagic manifestations but had a low platelet count (114 x 10^9/L). The patient's highest viral load was recorded on day 6 (cT 14.3), and the patient's symptoms severity roughly correlated with viral load until day 6. The patient was administered with ZMapp (Mapp Bioscience) on day 5 and day 8 [199]. ZMapp is a human chimaerised variant of ZMab, a cocktail of antibodies developed by Mapp Biosciences and the Public Health Agency of Canada. A detailed time course of viral FPKM (Fragments per kilobase per million reads) and viral load measured by qRT-PCR is displayed in Figure 3.3.

The patient viral load data was compared with diagnostic findings from the outbreak (3.2.2) where those with a cT of <25 were more likely to suffer fatal outcome, and those with a cT of >25 had CFR of <30%. Patients in the general outbreak with cT of ~14 invariably suffered fatal outcome. This illustrated the severity of the patient's illness and it was expected that the patient would not survive. After second treatment with ZMapp viral load decreased and was not detectable after day 12 of illness. This should not be interpreted as a result of ZMapp treatment however, as viraemia in non-fatal cases was observed to follow a similar pattern in other patients (200, 201). While a fatal outcome was expected the patient survived and convalesced [234]. On day 21 the patient had a pronounced thrombocytosis (peaking at 1726 x 10^9/L) as reported in Wilson et al 2016 [234]. The patient was tested by qRT-PCR further on days 25 and 28 of illness, and was discharged after two consecutive negative results, in accordance with general WHO advice. The overall course of illness was very similar to the clinical findings of outbreaks caused by other variants of EBOV [130].

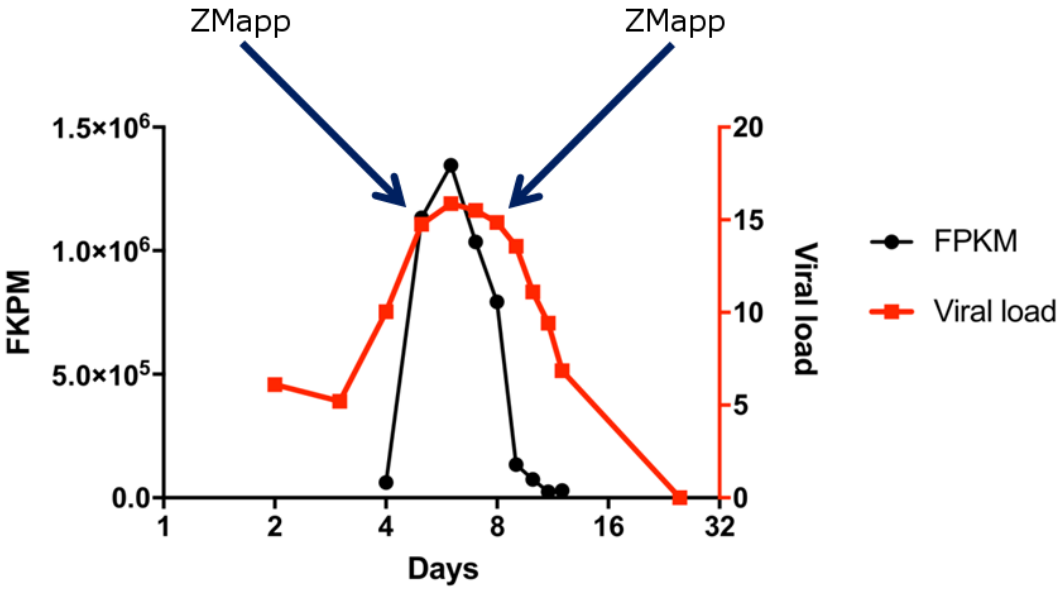


Figure 3.3: Time-course showing viral load as measured by qRT-PCR of the GP gene of EBOV in samples collected from a UK case of EVD. The right-hand y-axis displays the relative viral genome abundance (calculated by PCR maximum cT (30) - sample cT to calculate the proportional inverse value). The left-hand y-axis displays the FPKM of the GP gene measured by RNASeq, and correlates well with qRT-PCR data. The peak of viraemia is discernible at day 6. The X-axis displays the days post symptoms. Details of illness corresponding to this time course are published (199) and an analysis of patient clotting published separately [234].

3.4.2 Antiviral response associated transcripts appeared to increase in abundance early in infection.

Data from next generation sequencing was mapped to the human transcriptome, showing trimmed read counts of between 2 and 4 million reads per sample. This was considered low depth for transcriptomic analysis, so interpretation was conservative and focussed only on genes which appeared consistently throughout the time course. 773 genes with FPKM of more than 10 in the Cufflinks datasets and analysed using the MORPHEUS clustering tool (Broad Institute), and were displayed in Figure 3.4. Hierarchal clustering was performed using 1D Pearson clustering. Clustering based on the maximum transcript abundance resulted in discrete clusters emerging that appear specific to each time point. It was observed that transcripts early in infection which clustered together had similar functional annotation using Ingenuity Pathway Analysis, indicating these transcripts encoded proteins involved in the cellular antiviral response, clusters corresponding to these early transcripts are highlighted in figure 3.4.

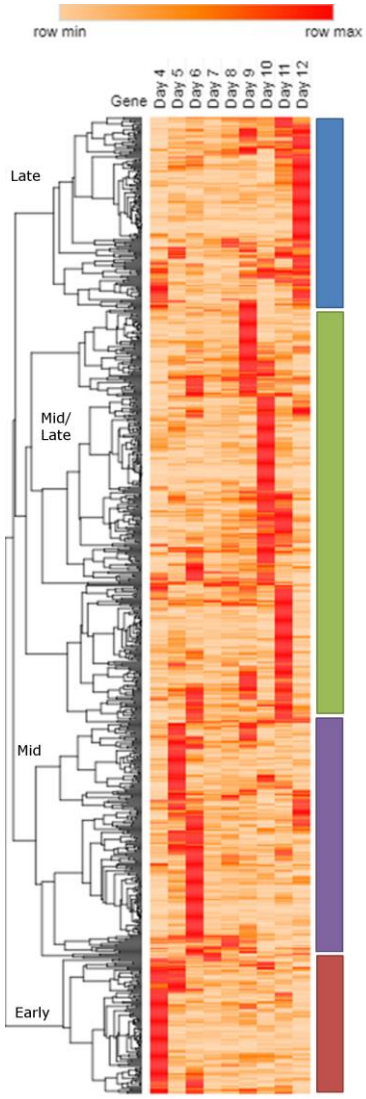


Figure 3.4: Transcriptomic analysis of UK clinical case of EVD using Cufflinks. FPKM values were analysed in MORPHEUS (Broad Institute) and displayed a heatmap with hierarchal clustering by row. Clusters of highest transcript abundance are visible for each day of the time course. These clusters have been interpreted to coincide with early (Red), Mid (Purple), Mid-Late (Green) and Late (Blue) stages of infection. Clustering was performed using the Pearsons clustering in a one dimensional analysis, only evaluating data for transcripts on each day sequentially. Clustering is also depicted as a dendrogram, which has been labelled Early, Mid, Mid/Late and Late to correspond with the clutsers annotated with coloured boxes. Colours are given on a global relative change for each row separately, given as row min ("Pale Yellow") to row max ("Red). Only transcripts detected in all 9 datasets are displayed (773 gene transcripts types)

3.4.3 Interferon associated transcript abundance correlated with viral load during severe EVD

Transcripts encoding well characterised interferon stimulated genes are shown in Figure 3.5. FPKM values for each transcript was compared at each time point and displayed without further processing. Patterns of transcript abundance for all 16 of these genes correlated well with viral load with high initial abundance, falling until day 12. Non-linear regression was performed on interferon gene transcript abundance compared with viral load over the 9 day time course, revealing that there was a slight correlation ~0.3-0.6 for the gene set ($p = <0.001$). For the small subset of transcripts shown in figure 3.5, the time point with the highest transcription level was similar to the highest level of viral genome abundance, though the calculated correlation was weak.

The significance of these changes is difficult discern without an appropriate age-matched and gender-matched control, therefore further analysis and looking for deeper meaning in the data is prohibited by these limitations. Clinical findings and viral load data clearly demonstrate that had this case not received effective therapeutic intervention, the outcome would almost certainly have been fatal. Contrasting clinical and molecular findings from this case suggest that the disease features are very similar to that described for EVD caused by other EBOV variants.

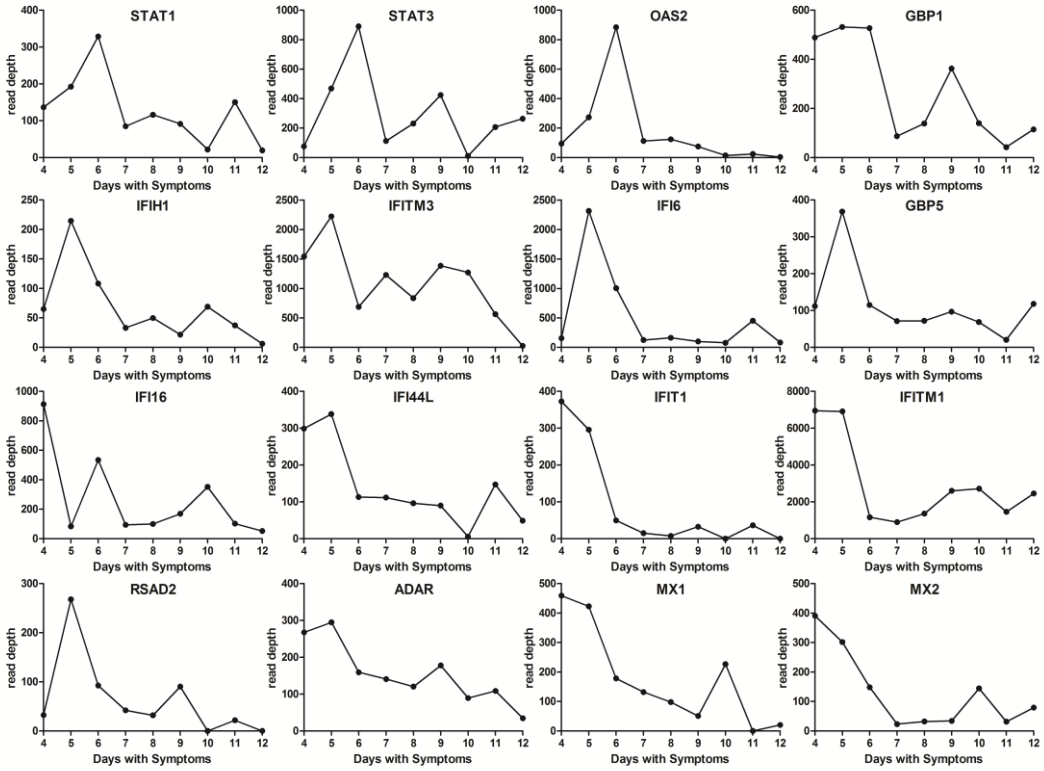


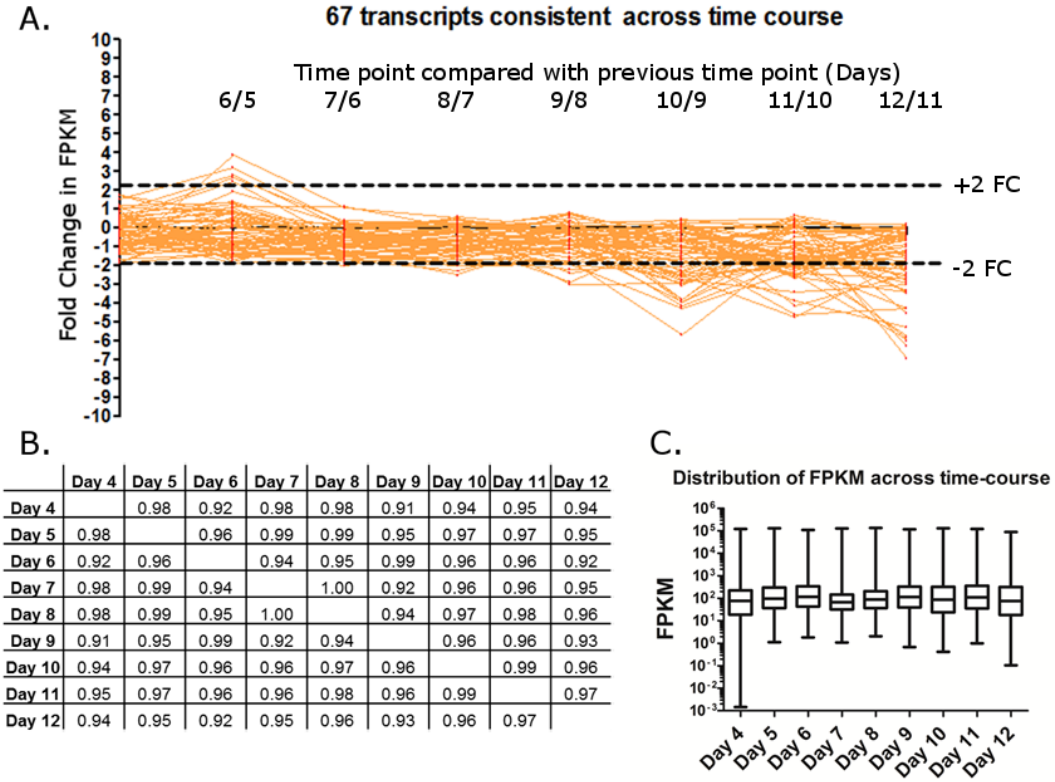
Figure 3.5: Transcript abundance data (FPKM) displayed for 16 genes characterised as known interferon stimulated genes (ISG). All genes were highest early in the time course, roughly similar to the observed levels of viral genome abundance and how it changes over time. Transcript FPKM is shown on the y-axis for each gene and x-axis displays the days with symptoms from day 4 to day 12 and titled with the term "read depth". The identifier for each gene is given above each chart. Sequencing data was acquired by Illumina MiSeq, and a single specimen collected by the clinical care team was sequenced twice in parallel, and read data combined in subsequent Cufflinks analysis as duplicate experiments, to enhance read depth.

3.4.4 Assessment of transcriptomic data does not indicate a bias due to high data diversity which may conflate transcript abundance changes.

As samples for this analysis were not prepared using optimum conditions, the abundance of transcripts was inspected across the time course to ensure that no major variation may introduce bias. The ranges of FPKM values were assessed to gauge the potential impact of size-factors, and a Spearman's correlation analysis was performed. Datasets correlated closely (>0.9 p=<0.001), a graphical representation of this analysis is given in figure 3.6. Overall there was no indication that the transcriptomic analysis may have been biased by data distribution anomalies.

3.4.5 Transcriptomic profiles suggest that infection related transcripts correlate with severest symptoms, but occur chronologically after the peak of viral load.

Transcripts with significantly changing abundance at each of the time-points (days 4-12), were aligned to KEGG pathways to assess functional categorisation. This analysis indicated that transcripts that were most significantly over-represented in this analysis aligned primarily to KEGG pathways involved in infection. Furthermore, activity scoring was performed based on the cumulative increase or decrease in transcript abundances in each category, indicating that over-represented transcripts showed the greatest activity at days 8-day 10, which occurs 2-4 days later than the peak of viraemia the patient experienced, but aligns with the period of severe illness reported by clinicians. It is therefore likely that these infections related KEGG pathways represent increased activity of the patient's immune system following the peak viraemic period. A graphical representation of this analysis is given in figure 3.7.



	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12
Day 4		0.98	0.92	0.98	0.98	0.91	0.94	0.95	0.94
Day 5	0.98		0.96	0.99	0.99	0.95	0.97	0.97	0.95
Day 6	0.92	0.96		0.94	0.95	0.99	0.96	0.96	0.92
Day 7	0.98	0.99	0.94		1.00	0.92	0.96	0.96	0.95
Day 8	0.98	0.99	0.95	1.00		0.94	0.97	0.98	0.96
Day 9	0.91	0.95	0.99	0.92	0.94		0.96	0.96	0.93
Day 10	0.94	0.97	0.96	0.96	0.97	0.96		0.99	0.96
Day 11	0.95	0.97	0.96	0.96	0.98	0.96	0.99		0.97
Day 12	0.94	0.95	0.92	0.95	0.96	0.93	0.96	0.97	

Figure 3.6: Data quality, ensuring data is not biased by size and distribution factors. (A) A selection of 67 transcripts with data at all time points, showing transcript abundance within 2 fold change in in FPKM at each day of the time course. Fold change was calculated by contrasting data on days 5-12 with data acquired from the previous data. (B) Correlation analysis of the datasets indicating that all datasets are >0.9, and there are no dramatic changes across the time course which may bias analysis. (C) Showing the distribution of FPKM across the time course, illustrating that the range in the data is similar at each time-point, and that size-factors are unlikely to bias the analysis. As only a single specimen was sequenced it was not appropriate to use statistical methods to calculate the relative dispersion of read depth across the sequenced information, instead these analyses attempt to describe the data and demonstrate any indications of data bias. It is notable that the Day 4 data has a much wider range extending into lower FPKM values, though the overall upper range and trimmed-mean is similar to other timepoints.

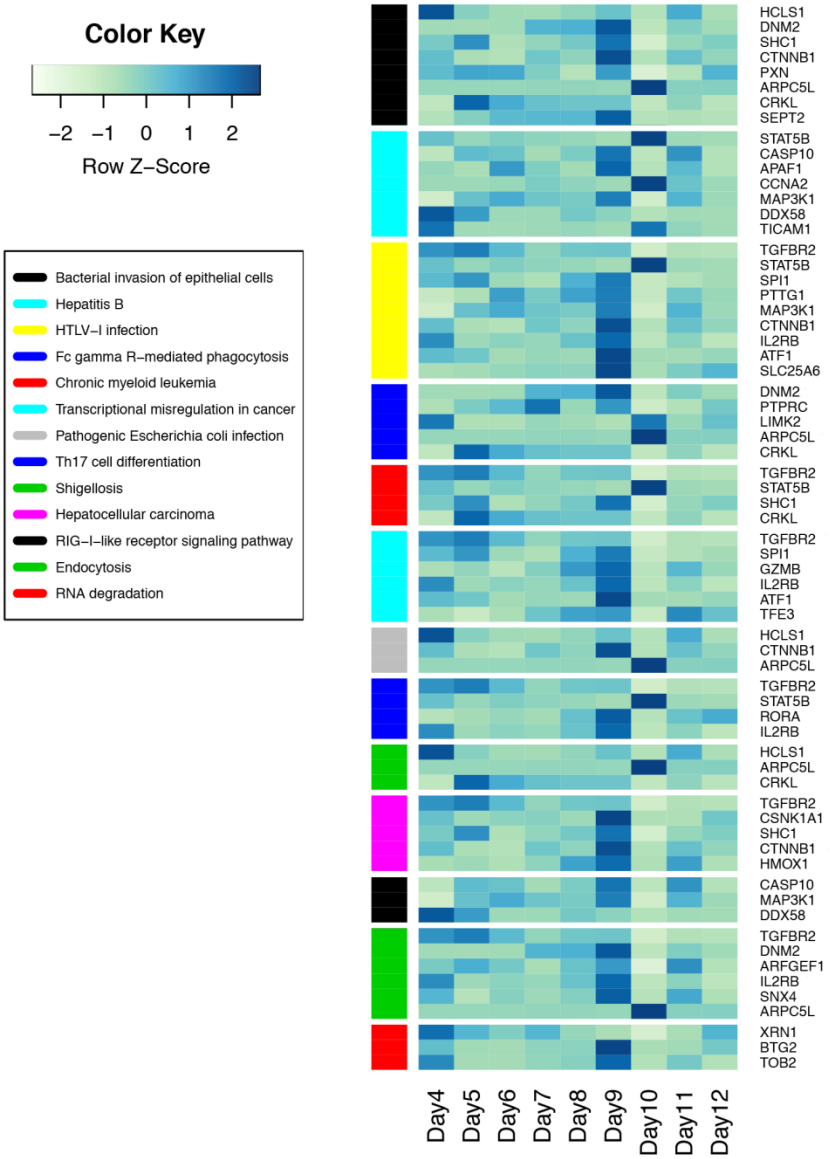


Figure 3.7: KEGG pathway analysis of transcripts which changed significantly at all time points. Transcripts were grouped together in hierarchal cluster analysis by ontology, by mapping genes to the KEGG database. The transcripts aligning to each KEGG pathway are indicated by colour, and the corresponding KEGG pathway is shown in the legend. Z-score (activity score) was calculated for each transcript at each time point, and is displayed in the clustergram ranging from z-score -2 (White) to z-score +2 (Dark blue).

3.5 The host response in A549 cells infected with the Makona variant was similar to infection with the Ecran variant

3.5.1 Viral RNA levels were similar between Makona and Ecran EBOV isolates

A549 cells were cultured to 80% confluence and infected at an MOI of 0.5 with Yambuku-Ecran or Makona-C05 stocks of virus. Infections were preformed at this low MOI to better profile the kinetics of viral culture, but likely resulted in reduced % of cells infected with virus as a result that was taken into consideration in the final analysis.

Analysis showed no significant difference in the abundance of viral genome in Makona and Ecran infected cells at 24h, 48h and 72h post infection (figure 5A). Therefore any changes observed in cellular transcription were not due to differing levels of viral replication. The abundance viral mRNAs was calculated by mapping reads to the EBOV genome, separating mRNAs by open-reading frame and normalising gene reads to viral gene length, shown in Figure 3.8B. There were no significant differences in viral gene transcript abundance between the Makona and Ecran variants. It was observed that patterns of transcriptional activity profiled by RNASeq in the case of both Makona and Ecran did not fit an expected gradient of transcription observed in other Mononegavirales. Of particular interest is the high level of GP transcript profiled in this experiment, which may indicate that the GP transcriptional products are more stable than other Ebola virus gene transcripts, resulting in a higher concentration of GP in infected cells. Overall these analyses indicated that the levels of virus and viral replication was similar in A549 cells for both Makona and Ecran.

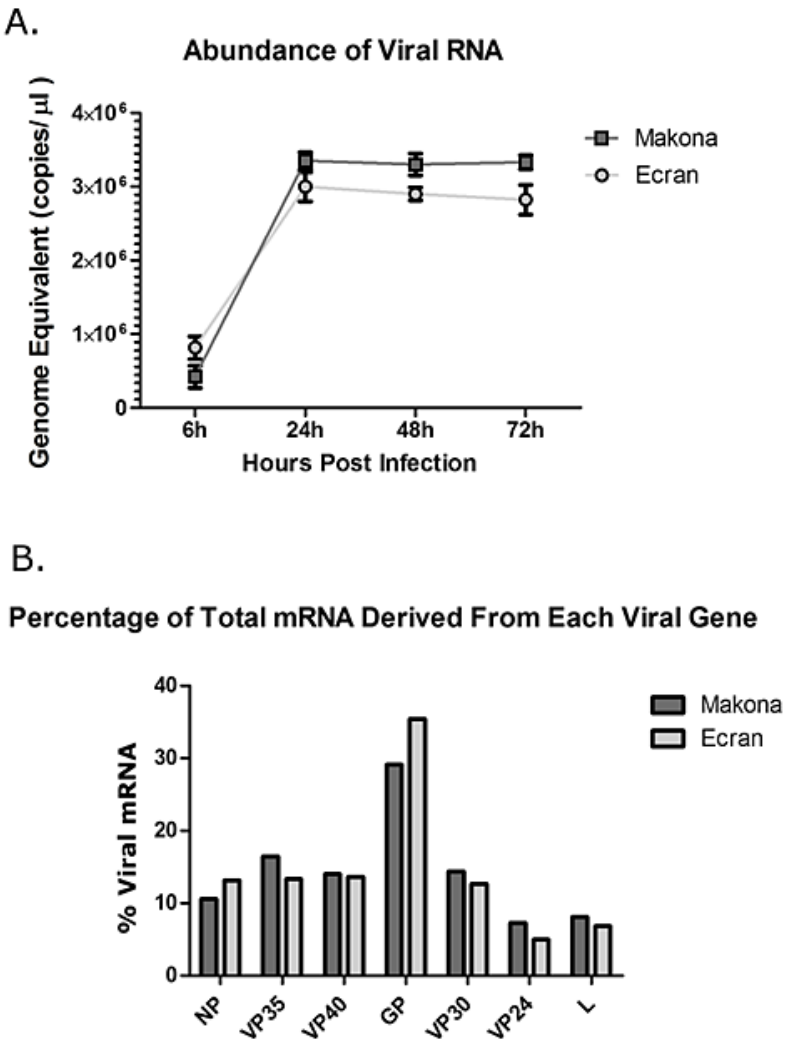


Figure 3.8: (A) qRT-PCR measurement of viral genome abundance in both Makona and Ecran infected cells (MOI 0.5). Measurements were performed on intracellular RNA, thus viral genome corresponds to intracellular virus which may be at multiple phases of viral lifecycle. Measurements are shown in genome equivalent copies/µl, by comparing qRT-PCR cT values to those of a known standard using synthetically created RNA. Shown are three assay technical replicates performed on biological duplicate samples at 6h, 24h, 48h and 72h. (B) Illumina next generation sequencing reads were mapped to viral transcript sequences, normalised to viral gene length and abundance measurements calculated. Values were normalised between viral variants by calculating the percentage (%) of reads mapping to transcripts corresponding to each viral gene. Shown on the x-axis are the symbols of each viral gene starting with the genomic 3' end (NP) and ending with the L gene at the 5' genomic end. For both (A) and (B) Makona is shown in dark grey and Ecran is shown in light grey.

3.5.2 Transcript abundance identified significant differences between genes

The total number of trimmed reads which met quality thresholds are shown in Table 3.9A, this total ranges from 57588535 at 48h in Makona infected A549 cells to 102746478 reads at 24h in Makona infected A549 cells. Those genes which showed a significant increase or decrease in transcript abundance are shown in Table 3.9B. This also shows the total number of genes in the database to which reads mapped.

Low numbers of reads can bias analyses by indicating a significant fold change when in reality this represents small transcript abundance and can be affected considerably by biological variations between samples. Conversely, a small fold change which corresponds to a change in transcript abundance numbering thousands of counts, could be considered a significant finding, and ascertaining whether biological variation would have a biasing effect at transcripts with greater CPM was important. To confidently assess the impact of biological variation between samples, a dispersion calculation was performed to calculate a CPM threshold, below which the significance of change cannot be reliably determined.

A.

Sample	Trimmed Reads
ECRAN_24h	77610333
ECRAN_48h	77211979
ECRAN_72h	67874000
MAKONA_24h	102746478
MAKONA_48h	57588535
MAKONA_72h	66319784

B.

	ECRAN			MAKONA		
	24h	48h	72h	24h	48h	72h
Total Genes with mapped reads	26621	27146	26395	27895	25825	26541
Differentially Expressed Genes	11557	12865	13021	10456	12328	13239
Not Differentially Expressed Genes	15064	14281	13374	17439	13497	13302
Genes which increased in transcript abundance	6093	7119	7319	5658	6674	7506
Genes which decreased in transcript abundance	5464	5746	5702	4798	5654	5733

Table 3.9: The tables above describe the RNASeq data acquired from infection of A549 cells with either the EBOV Makona variant or the Ecran variant. RNA was collected at 24h, 48h and 72h, and each corresponding dataset is represented separately. (A) This table shows the total reads which were trimmed and met quality specifications, and were then used to produce read counts of transcripts. (B) This table shows the number of genes which significantly increased or decreased in transcript abundance across the experiment, as well as the number of genes which did not.

3.5.3 Large numbers of transcripts with abundance changes were identified when comparing Makona or Ecran infection with mock infection.

Tag-wise dispersion was calculated using EdgeR and plotted against logCPM using data from 3 biological replicates of mock infection, this is shown in figure 3.9A. This analysis revealed that transcripts with a normalised logCPM of less than -1 suffered a large degree of variation between these replicate mock infected samples.

Dispersion was used to more accurately identify genes with significantly changing transcript abundance. Dispersion is a measure of biological and technical variation in a dataset, used to compensate for potential biased diversity of data at different abundances measured, or in different specimens. Dispersion analysis demonstrated that transcripts with extremely low abundance (Log2 counts per million of less than <0) and those with high abundance that occurred infrequently in the dataset, had raised diversity, and should be accounted for when determining significant changes. Fold change data for Makona or Ecran infection of A549 cells at 24h, 48h or 72h compared with mock infection was plotted against the log2CPM for corresponding genes, This analysis is shown in figure 3.9B, and is a visual representation of data shown in Table 3.9B. An additional analysis was performed; plotting fold change comparing Makona infection directly with Ecran infection, to identify genes with significant differences in transcript abundance between Makona and Ecran infection of A549 cells. This analysis was plotted in figure 3.9C and illustrates the small numbers of significant transcript abundance changes comparing Makona with Ecran directly.

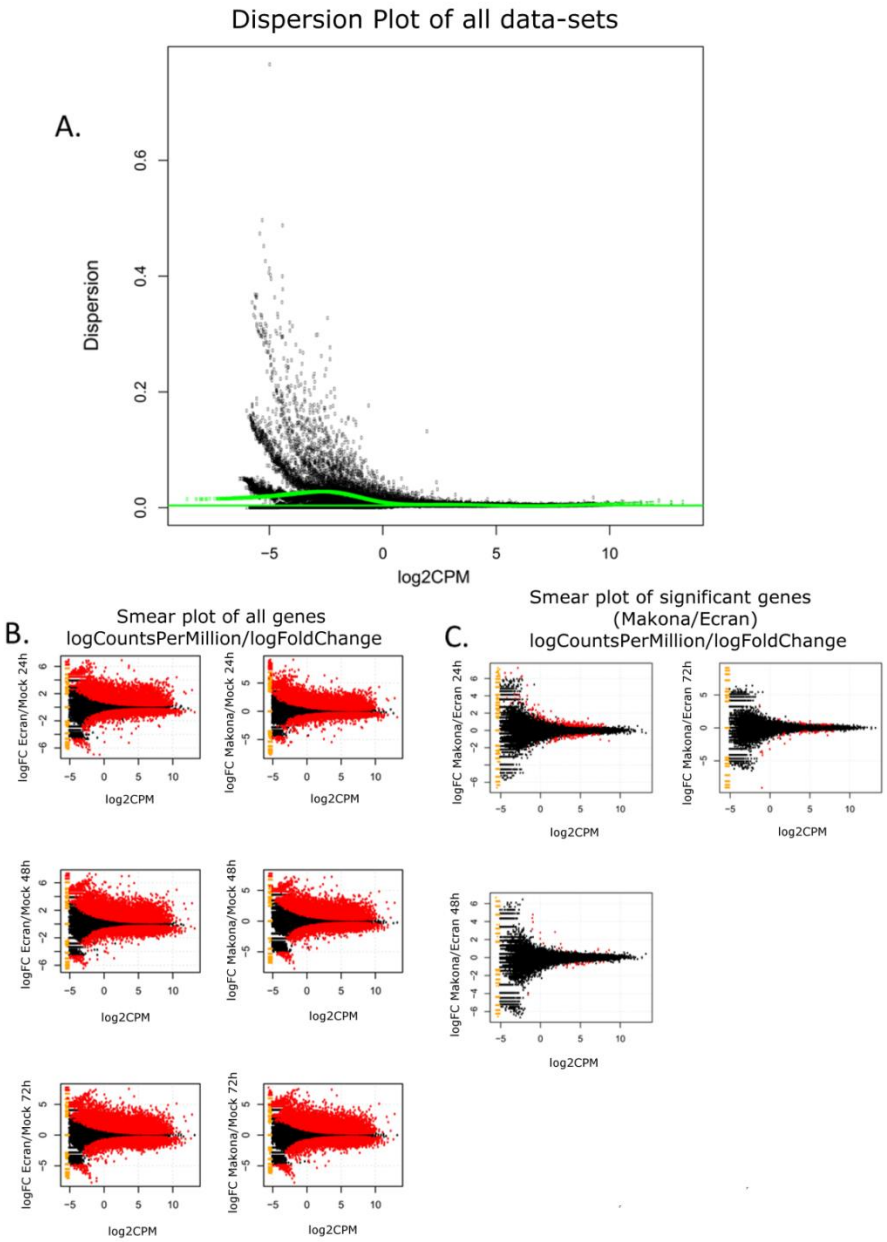


Figure 3.10: (A) A dispersion plot showing tag-wise dispersion, plotting estimated dispersion against logCPM in data collected from three independent biological replicate samples of mock infection. (B) Smear plots showing log2CPM against fold change, accounting for the effect of dispersion to identify genes with significant changes in transcript abundance. These plots show significant genes in red and insignificant genes in black, comparing Ecran and Makona infection with mock infection at each time point. (C) Smear plots showing genes with significant changes in transcript abundance in red, and those without significant changes in black comparing Makona infection with Ecran infection directly at each time point.

3.5.4 Data modelling showed closest correlation between Makona and Ecran induced host responses at 72h post-infection.

A principal component analysis was performed to assess the correlation between transcript abundance data in Ecran and Makona infection of A549 cells at 24h, 48h, 72h and mock infected A549 cells. Principal component analysis was described in detail in Chapter 2, subsection 2.7.5. Transcript counts were orthogonally transformed into linear uncorrelated values representing the data as a whole. These values were then plotted in a 2D space represented in Figure 3.11. The distance between each variable is indirectly proportional to how well they correlate. This analysis reveals that transcript abundance data from infected A549 cells does not correlate well with mock infected cells, as one would expect. Furthermore, the results of the PCA reveals that the Ecran and Makona variants induce cell transcript abundance changes which were very different at 24h post infection, and most similar at 72h post infection.

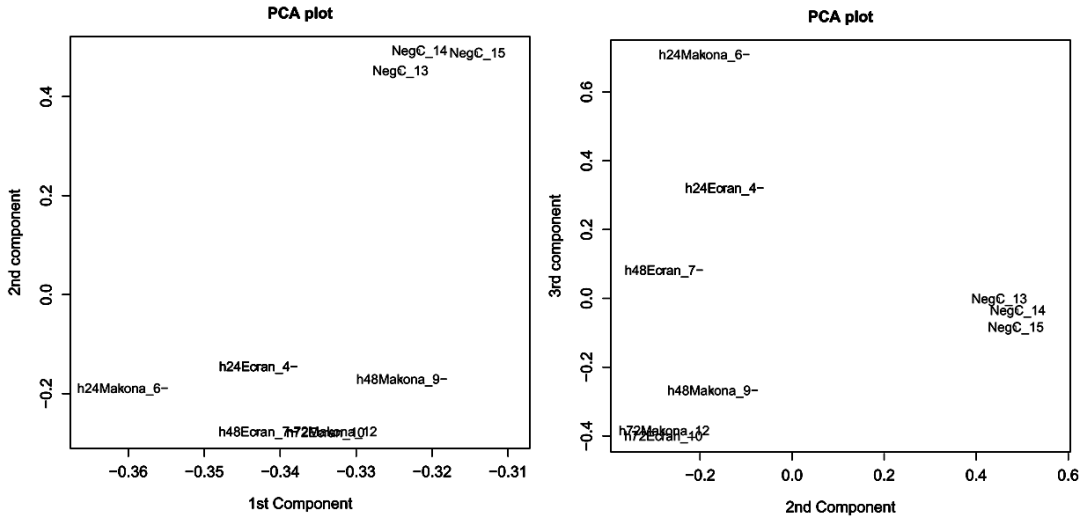


Figure 3.11: Principal Component Analysis (PCA). Transcript abundance data was converted using an orthogonal transformation to identify the presence of a correlation between variables. PCA plot shows the results of this analysis in a 2D space. The proximity of each variable is proportional to the similarity of the data. It can be clearly seen that mock infected data (NegC) clusters closely together, and can be separated by both the 1st and 2nd components and the 2nd and 3rd components from samples derived from infected A549 cells. Additionally, it can be observed that by 72h the Ecran and Makona variables correlate very closely, but they correlate least well at 24h.

3.5.5 The host response to Makona and Ecran were most different at 24h post infection.

Transcript abundance data was compared between Makona or Ecran infected A549 cells at 24h, 48h or 72h with mock infection. Log2 fold change was calculated and was displayed as a heatmap in Figure 3.12. Hierarchal clustering was performed on transcripts (rows) only, forming 15 discrete clusters. At a transcript level, the log2 fold change data looks very similar in all datasets. To better assess for correlation at the resolution of individual transcripts, a second heatmap was produced showing log2 fold change in transcript abundance in Ecran infected A549 cells directly compared with Makona infected A549 cells at 24h, 48h and 72h. This heatmap was shown in Figure 3.13. Hierarchal clustering was performed and transcripts clustered into 15 discrete clusters. The greatest log2 fold change in transcript abundances occurred at 24h comparing Ecran with Makona infection directly. Combined with principal component analysis, this data indicated that the host response to infection with the Ecran variant or the Makona variant were most different at 24h. Later in infection, the data showed close correlation, indicating that by 72h the dynamics of late phase host response were much the same in A549 cells infected with these EBOV variants.

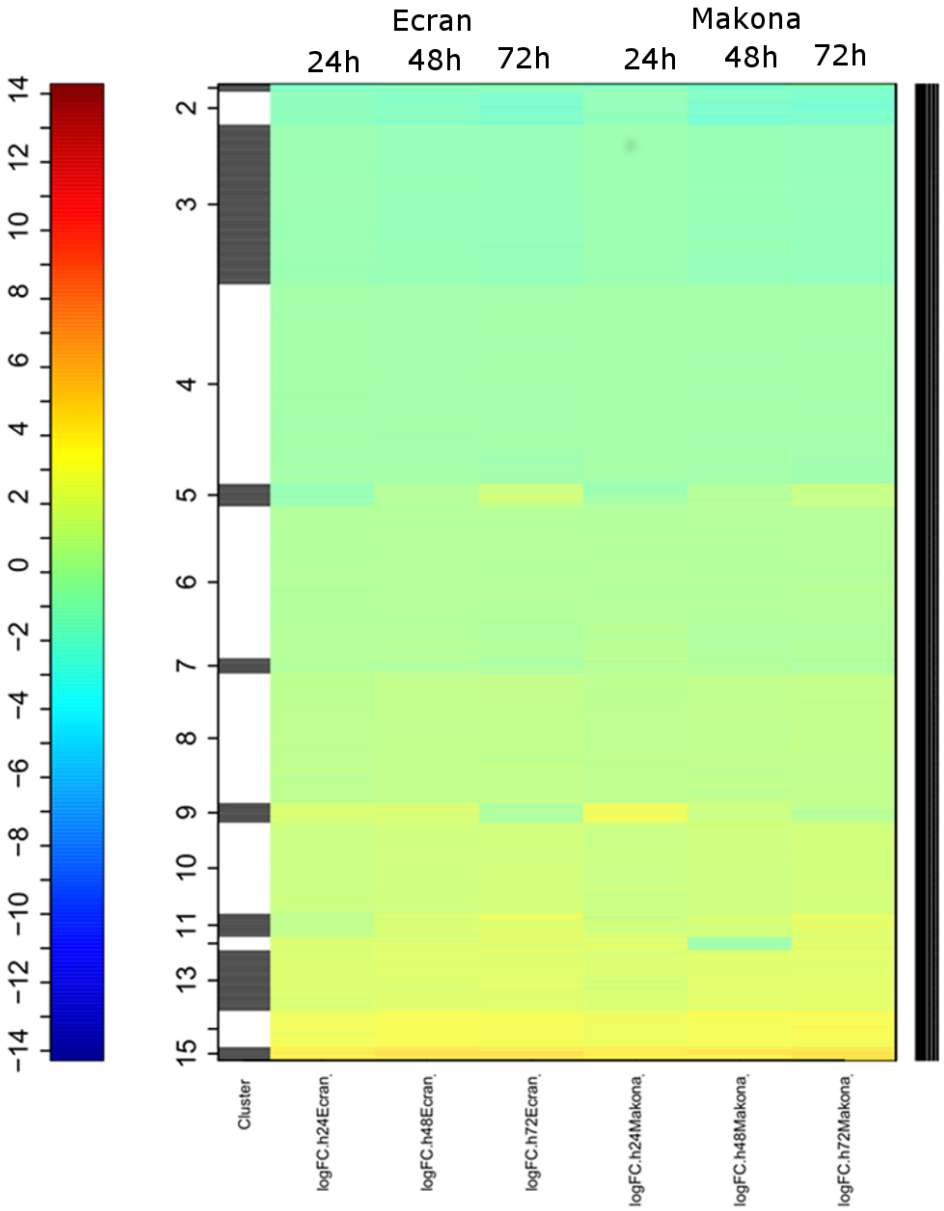


Figure 3.12: Heatmap comparing datasets derived from infected A549 cells with mock infection. This heatmap shows all transcripts displaying a significant change in abundance comparing Ecran and Makona infection at 24h, 48h and 72h with mock infection. The heatmap colours are scaled from 14 log2 fold change (Red) to -14 log2 fold change (Blue). One dimensional hierarchal clustering was performed using Spearman coefficient on transcripts with data forming 15 clusters by similarity in expression.

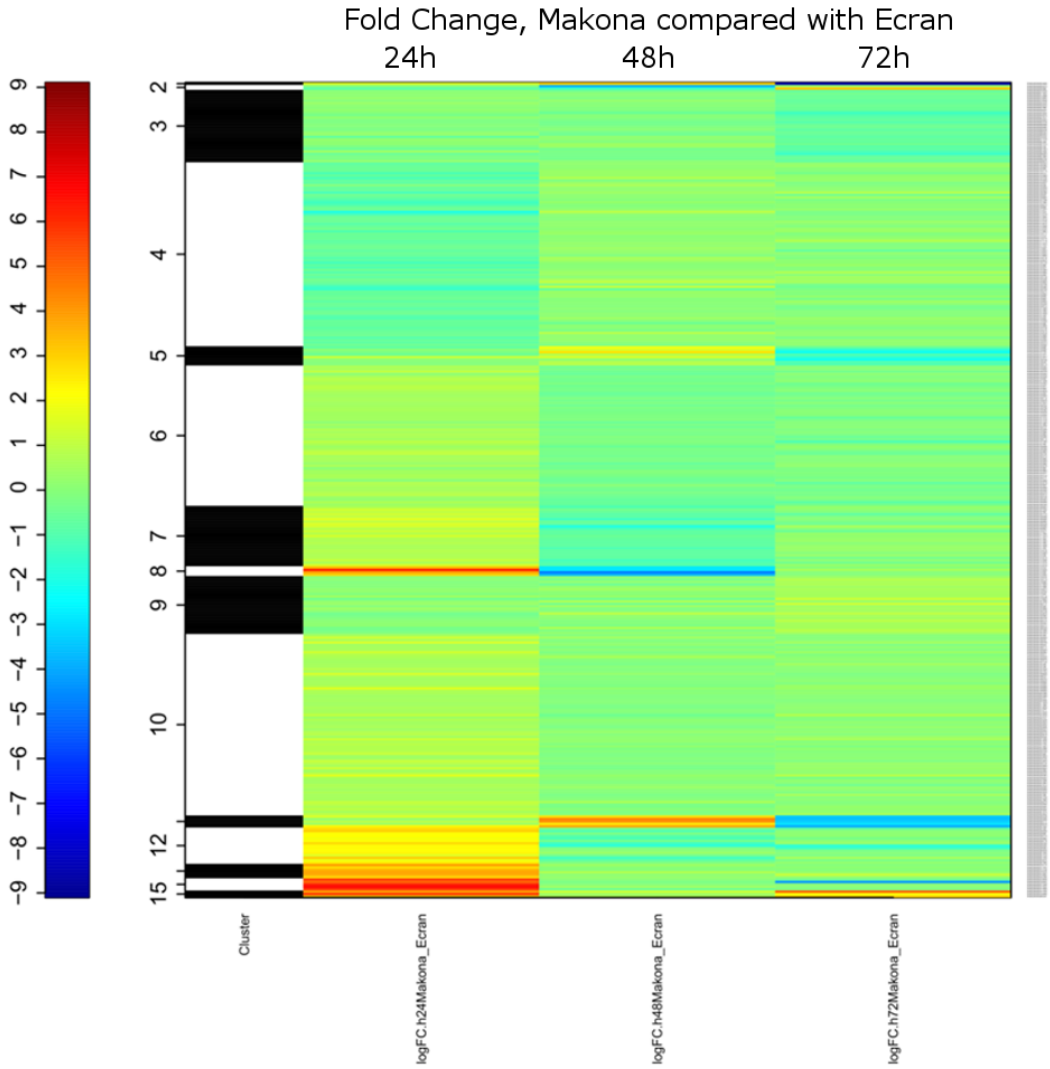


Figure 3.13: Heatmap showing transcript abundance data derived from Ecran infected A549 cells at 24h, 48h and 72h compared directly with Makona infected A549 cells. Only transcripts which displayed significant changes in abundance are shown. This heatmap shows a colour scale from 9 log2 fold change (Red) to -9 log2 fold change (Blue). Hierarchal clustering was performed one dimensionally, only on transcript (rows). This analysis showed that transcripts formed 15 discrete clusters, according to patterns of expression. It was clear from this analysis that the greatest log2 fold changes in transcript abundances occurred at 24h, and that log2 fold changes were not as great at 48h and 72h.

3.5.6 Transcripts associated with viral infection showed similar abundance changes in both Makona and Ecran infection.

Transcripts previously found to be associated with viral infection were investigated for significant abundance changes. Transcript abundance measurements calculated to be below the false discovery rate threshold (FDR), were visualised by heatmap, shown in Figure 3.14. Only genes exceeding a logCPM value of 5, with an FDR of <0.05 are shown. No threshold for fold change was set to allow a complete overview of transcript abundance at 24h, 48h and 72h time points for both Ecran and Makona infections. Gene Ontology was summarised on the left of the heatmap as interferon stimulated gene transcripts, stress associated transcripts, chemokine encoding transcripts, interleukin encoding transcripts and TNF associated transcripts. A histogram was provided on the scale bar in blue to illustrate the number of transcripts which fall into a particular segment of the colour scale.

ISGs IFIT3, IFIT2 and IFITM10 show increased log2FC at 48h and 72h compared with 24h post infection for both viruses. Transcripts encoding FADD (Fas-associated death domain) appeared significantly elevated in Ecran infection after 24h compared with mock infection, and was also raised in comparison to Makona at the same time point. Transcripts encoding Interleukins (IL15, IL11, IL27RA, IL6R, IL7R, IL20RB) were significantly raised in Ecran infection at 24h compared with both Makona infected and mock infected cells. IL15 is a potent inflammatory cytokine and appeared elevated in at all time points for both viruses, but was most acutely elevated at 24h post infection with Ecran. TNF signalling associated transcripts were also visualised, showing similar patterns. TNFAIP6 transcripts appear elevated at 24h in Ecran, and 48h in both Ecran and Makona.

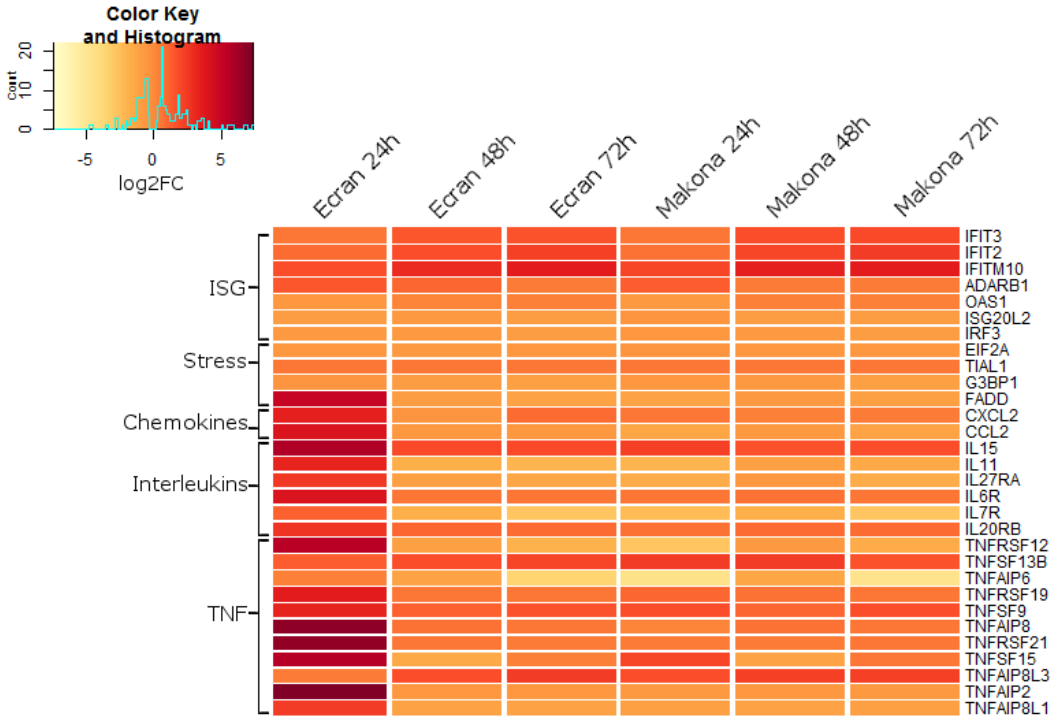


Figure 3.14: Heatmap showing log2FC in transcript abundance of genes in 5 categories compared with mock infected cells. Categories shown are Interferon stimulated genes (ISG), Stress response genes (Stress), Chemokines, Interleukins and TNF associated genes (TNF). A colour scale bar is provided to aid interpretation showing change in colour from -7 to +7 with threshold values set at -5 ("Pale Yellow") and +5 ("Blood Red") for minimal and maximal colours. 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale.

3.5.7 Transcripts involved in immune signalling showed similar abundance in both Makona and Ecran infection.

Genes with expression induced by EBOV infection in patients and animals were also investigated (143, 202, 203). Transcripts associated with the immune response, those encoding immune signalling proteins, those associated with Toll-Like receptor signalling (TLR) and those encoding metalloproteases were visualised by heatmap. The heatmap was generated as discussed in section 3.5.4. This analysis included genes which were calculated to be significant at least one time point. A histogram in the scale bar was provided to indicate the number of transcripts exhibiting log2FC values falling within a particular segment of the coloured scale.

Transcripts encoding STAT1, STAT2, STAT3, STAT4 and STAT6 are shown categorised under immune response. Abundance changes in these transcripts were usually observed in cells affected by a cellular immune response, and a good indicator of the strength of immune signalling activity. STAT1 appeared to remain similar to mock infected cells throughout the time course for both viruses. STAT2 appeared slightly elevated at 48h and 72h for both viruses, and correlates with downstream activity observed in ISGs shown in Figure 3.14. Transcripts encoding IFNA1 were also examined, showing a significant fall in abundance at 72h post infection with Makona compared with both Ecran and mock infection.

Immune signalling genes (CD40, CD80, IL6 and CXCL8 (IL7)) are also shown. CD80 is a co-receptor for T-cell activation; transcript abundance in Ecran infected cells appeared reduced at 24h and 48h compared with mock infected cells, but raised in Makona infection by 72h. IL6 has been associated with poor human patient outcome in EBOV infection (124), and appeared increased by 24h post infection in both Ecran and Makona. TLR associated

signalling transcripts, including those encoding subunits of the NFκB complex are shown. Most of these transcripts appear similar across all time points in both Ecran and Makona infection. NFKBIZ however appeared significantly raised at all time points, and was greatest at 72h post infection. NFκB signalling is a secondary phase response and is usually viewed as a late response to infection, and is readily induced by TLR4 signalling triggered by lipopolysaccharide (LPS). NFKBIZ encodes a gene product of the nuclear I kappa B family, believed to be involved in the regulation of IL6 production.

Metalloproteases were examined as they have been previously associated with lethal outcome in mouse models of EBOV infection (143). Metalloproteases are involved in tissue remodelling and affecting structural changes on the extracellular matrix. MMP9 is involved in numerous physiological processes including tissue remodelling by breaking down collagen.

MMP9 is also involved in angiogenesis and mediates neutrophil/leukocyte migration and extravasation. In this experiment transcripts encoding MMP9 appear significantly reduced in abundance at 48h and 72h post infection with both Makona and Ecran. Other transcripts encoding metalloproteases were also raised in abundance, including MMP13 which was found to be associated with lethal outcome in the mouse model of EBOV. MMP16 and MMP26 were also raised in abundance at all time points. MMP26 may be significant as this enzyme is involved in the cleavage of MMP9, and degradation of coagulation proteins including fibronectin and fibrinogen. Fibronectin (FN1) transcription increased overtime in both Ecran infection (0.36 log2FC at 24h to 1.43 log2FC at 72h) and Makona infection (0.38 log2FC at 24h to 1.15 log2FC at 72h). The abundance of transcripts encoding Fibrinogen (FGA, FGA, FGB) also increased overtime in both Ecran (average -4.23 log2FC at 24h to 0.29 log2FC at 72h) and Makona infection (-4.48 log2FC at 24h to 0.57 log2FC at 72h),

with a major drop in expression at 24h post infection compared with mock infection. Other coagulation associated transcripts including those encoding F2 and F3 (Tissue factor) were stable, in concordance with previous observations of these genes in EBOV infection. At 24h F2 (average 1.1 log2FC in Ecran and Makona) displayed a significant increase above mock infection and decreased by >2 fold to an average of -0.7 by 72h. F3 showed an insignificant decrease of -0.5 at 24h, decreasing further to -1.3 in Makona and -0.9 in Ecran.

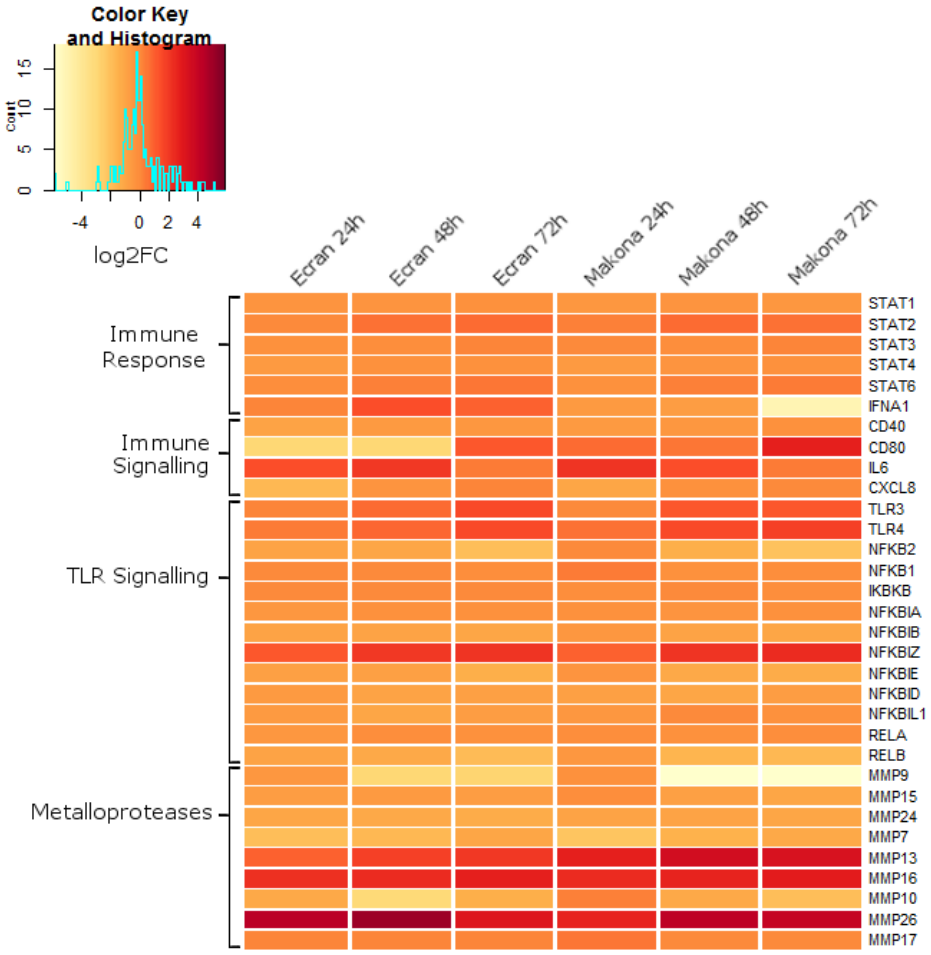


Figure 3.15: Heatmap showing log2FC in transcript abundance of genes in 4 categories compared with mock infected cells. Categories shown are immune response transcripts, those involved in immune signalling, TLR signalling associated transcripts and transcripts encoding metalloproteases. A colour scale bar is provided to aid interpretation showing change in colour from -6 to +6 with threshold values set at -4 ("Pale Yellow") and +4 ("Blood Red") for minimal and maximal colours. 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale.

3.5.8 Type I and II interferon signalling were similar in Makona and Ecran infection.

Patterns of interferon signalling were similar in A549 cells infected with both Makona and Ecran, suggesting these viruses antagonise/stimulate interferon with a similar efficacy. Considering this in addition to evidence of minimal genetic variation in the main interferon antagonist proteins (VP24 and VP35), suggesting that Makona was just as capable of limiting the interferon response to viral infection as Ecran.

To further analyse patterns of interferon signalling in Makona infected A549 cells, transcripts with significant abundance changes (more than +2 or less than -2 fold change, and <0.05 FDR) were selected for analysis using the Interferome v2.1 database. This database of interferon signalling associated genes helped identify interferon stimulated/suppressed transcripts in the transcriptomic data. 89 significantly changing transcripts matched those in the interferome database. 38 were involved in type I interferon signalling, 34 in type II, and 3 were involved in type III. An additional 40 were involved in both type I and II signalling and 5 appeared related to all three types of interferon signalling. No identified transcripts were significantly changed in Makona infection compared with Ecran. However this analysis indicated a significant number of genes appeared changed in Makona infection compared with mock infected A549 cells.

3.5.9 Canonical pathway activity appeared increased

Transcript abundance measurements were analysed with Ingenuity Pathway Analysis to test the biological significance and potential impact of abundance changes on cell biology. Activation scores (z-scores) were calculated and the predicted activity of 135 pathways were assessed. >50% of pathways increased in z-score during infection with either Ecran or Makona and most pathways displayed similar activity in response to both viruses. Pathways include those related to immune response, NFκB signalling, NFAT signalling, macrophage nitric oxide and reactive oxygen species signalling, MAPK signalling, Rho Family GTPase signalling, CD40, IL-8, and IL-3 signalling pathways.

Apoptosis related pathways including the PI3K/Akt, TGFβand "Apoptosis induced by HIV1 infection" pathways; which include genes involved in the intrinsic pathway of apoptosis, showed a reduced activity. These patterns were in concordance with previous literature (136, 143, 204, 205), and studies finding that EBOV does not induce apoptosis (206).

The EIF2 pathway activity decreased in activation score, The abundance of transcript transcribed from the EIF2A and EIF2B genes does not appear similarly affected, through EIF2B3 appeared significantly reduced in abundance (average -1.2 log2FC in both Ecran and Makona) at 48h and 72h post infection. PERK (EIF2AK3) and GCN2 (EIF2AK1) phosphorylate EIF2A in response to ER stress and nutrient depravation respectively, but these transcripts did not significantly change in abundance. The type 1 interferon response is also known to modulate EIF2 activity. EIF2 is a translational suppressor activated during cell stress, and mediator of stress granule formation (167).

3.5.10 Cholesterol Biosynthesis appeared affected by EBOV infection

Pathway analysis also revealed the top 10 pathways to which significantly changing transcripts aligned, shown in Table 3.16. The top pathways were several cholesterol biosynthesis pathways. EBOV has been shown to use proteins involved in cholesterol pathways to enter and infect cells, namely NPC1 (85, 93), which showed a change in transcript abundance (log2 2.265 and 1.847 increase for Makona and Ecran, respectively) and an upstream activator known as transcription factor 7 like 2 (TCF7L2), which increased in log2 fold abundance in Makona and Ecran infected cells by 2.789 and 2.819, respectively.

Ingenuity Canonical Pathways	-log(p-value)	z-score	Molecules
Superpathway of Cholesterol Biosynthesis	7.85E+00	NaN	MVD,SQLE,ACAT2,IDI1,MVK,HSD17B7,MSMO1,TM7SF2,SC5D,FDPS,FDFT1,DHCR7,DHCR24,LSS,HMGCR,HMGCS1
Cholesterol Biosynthesis I	5.61E+00	NaN	SQLE,FDFT1,DHCR7,DHCR24,HSD17B7,MSMO1,LSS,TM7SF2,SC5D
Cholesterol Biosynthesis II (via 24,25-dihydrolanosterol)	5.61E+00	NaN	SQLE,FDFT1,DHCR7,DHCR24,HSD17B7,MSMO1,LSS,TM7SF2,SC5D
Cholesterol Biosynthesis III (via Desmosterol)	5.61E+00	NaN	SQLE,FDFT1,DHCR7,DHCR24,HSD17B7,MSMO1,LSS,TM7SF2,SC5D
Amyotrophic Lateral Sclerosis Signaling	3.71E+00	NaN	NOS1,GRID2,GRID1,GPX1,PIK3R5,PDGFC,PGF,BCL2,GRIK5,CACNA1E,HECW1,FIGF,AKT3,GRIK2,PPP3CA,GRIK1,GRIN2B,CACNA1D,NEFM,CACNA1C,VEGFC,CACNA1A,CAPN8,GRIK4,GRIA3
G-Protein Coupled Receptor Signaling	3.68E+00	NaN	HTR2B,VIPR2,ADCY4,CAMK2A,PDE7B,HTR7,RGS14,PDE11A,MC2R,ADRA1B,PDE2A,PDE10A,GRM1,GRM8,CREBBP,GRM4,PDE4B,CREB5,CHRM5,GRM7,GNAO1,ADCY10,FYN,GRM2,PIK3R5,GNA14,RAPGEF4,CHRM3,PRKAG1,EP300,BRAF,SYNGAP1,DRD1,PDE3B,AKT3,PRKCE,PDE4D,DRD4,ADCY8,ADORA1,PRKCA,RGS2,RGS7,GNAQ,PRKCG,GABBR2,PLCB4,P2RY14,CALCR,HTR1F,PDE8B
Xenobiotic Metabolism Signaling	3.55E+00	NaN	CYP3A7,CAMK1D,MAF,MAPK13,IL6,CYP2C19,UGT2B15,CHST2,MAOB,ALDH1L1,ALDH1A1,CAMK2A,PPM1L,CHST11,HS6ST3,ALDH3A1,PRKD1,MGST1,HDAC4,CREBBP,ALDH8A1,GRIP1,HDAC5,HS3ST3B1,PPP2R3A,ALDH1A2,PPP2R2B,NCOA1,PRKCH,SLCO1B1,MAP2K5,GAL3ST2,HS2ST1,PIK3R5,FMO5,EP300,HMOX1,GSTT2/GSTT2B,UST,HS6ST2,PRKCE,PPP2R2C,NOS2,NFE2L2,PRKCA,CYP1A1,SULT1C2,MGMT,MAPK8,PRKCG,SULT2B1,AHRR,SULT1B1
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	3.40E+00	NaN	MVD,FDPS,ACAT2,IDI1,COX10,MVK,HMGCR,HMGCS1
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	3.13E+00	NaN	MMP3,LTB,IL6,IL1R2,CAMK2A,DKK2,FIGF,GSK3B,PLCL1,IL1RAP,TNFSF13B,PRKD1,TNFRSF11B,WNT9A,CREBBP,VEGFC,PLCL2,CREB5,PLCG2,GNAO1,PRKCH,IL1RAPL1,PDGFD,TCF4,FRZB,PIK3R5,PDGFC,PGF,EP300,ROR2,WNT7A,NFAT5,AKT3,SFRP5,PRKCE,NOS2,PPP3CA,MMP1,PRKCA,C5AR1,DAAM1,IL15,GNAQ,TCF7L1,NFATC4,PRKCG,FOS,PLCB4,CSF1,LTA,DKK1,TCF7L2,LRP1,WNT11,IRAK2
Factors Promoting Cardiogenesis in Vertebrates	2.93E+00	NaN	NOX4,TCF4,BMP4,BMP8A,TGFBR3,ACVR1,TCF7L1,BMPR1B,PRKCG,PRKCE,MEF2C,PRKCH,BMP7,DKK1,GSK3B,BMP6,ACVR1C,WNT11,LRP1,TCF7L2,PRKD1,PRKCA

Table 3.16: Gene transcripts showing a significant change in abundance in Makona infected A549 cells 24h post infection compared with mock infected cells were aligned to canonical pathways. This table shows the top10 most significant pathways identified in this analysis, and indicated the functional pathways to which these genes align. –log10 p-value is calculated to show the significance of the alignment of genes. The table is arranged from the most significant pathway to the least significant pathway of the top 10. The gene transcripts aligning to that pathway identified with RNASeq analysis are shown in the 4^{th} column.

3.5.11 Upstream regulators involved in cell survival, stress and immune activation show some differences in activity

Less than 1% of transcripts identified in this study were different in Makona compared with Ecran infected cells. Transcripts which changed >±1 log2FC and <0.05 FDR were selected for canonical pathway analysis. This analysis was used to determine significant upstream regulators.

The top 10 most significant pathways identified were shown in Figure 3.17 (left), and show that TNFR2 Signalling shows the highest predictive value. TNFR2 is a surface receptor of TNF, which activates downstream regulators including NFκB, MAPK and AP-1, all of which were involved in cell activation and the host immune response.

MAPK transcript abundance did not change significantly over the course of infection with either virus. Transcripts encoding the AP-1 components FOS and JUN however did change significantly, with FOS increasing over time from 2 fold change to over 8 fold change when infected with either virus compared with mock infection. JUN however decreased overtime by ~2 fold change throughout infection. NFκB associated transcripts abundances were the same in both Makona and Ecran infection.

Upstream regulator prediction was used to identify common regulatory proteins which may evoke the small differences seen in the host response to Makona compared with Ecran. The top 10 most significant predicted upstream regulators were shown in Figure 3.17 (right). These include the NFκB complex (including RELA), TNF associated regulators (TNF and TRADD), the pro-inflammatory IL1ββ and the stress response mediator EIF2AK2. Functional networks to which these regulators align include cell survival, immune activation and cell stress.

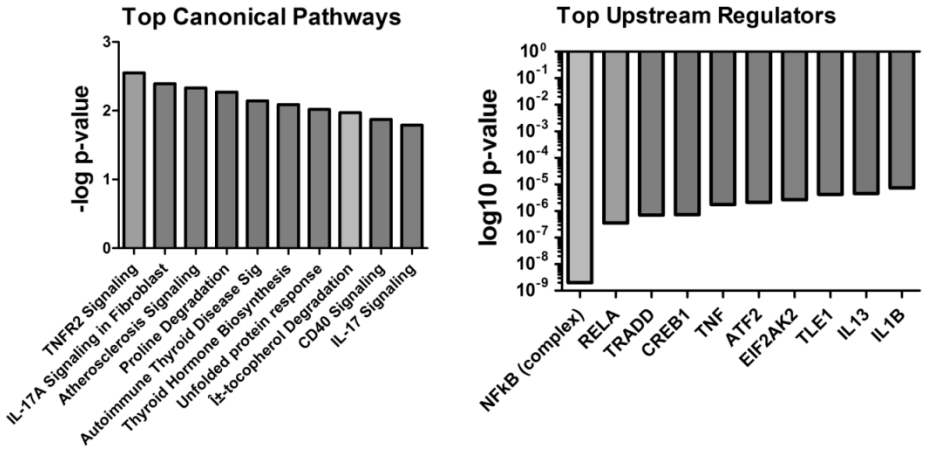


Figure 3.17: Genes identified as significantly different between EBOV Makona and Ecran were analysed in ingenuity pathway analysis. Upstream regulators and transcription factors with associations common to significantly changing transcripts were identified. The top 10 most significant upstream regulators were shown. Log10 p-value indicated the significance of the predicted upstream regulators, with lower values indicating more significance.

3.5.12 Transcripts showing significantly different abundance were associated with the inflammatory response

Transcripts which were under the regulatory power of the top 10 significant upstream regulators were investigated. A network map was generated by computing experimentally verified direct and indirect interactions. 21 genes were identified in a sub-network and all orphaned genes without connections were removed from the analysis, this network is shown in Figure 3.18. Network connections were confirmed using STRING 10 to ensure interactions were shared across databases. Functional analysis was performed in IPA, 4 transcripts mapping confidently to Leukocyte Extravasation, 4 mapping to Proliferation of Cells and 13 genes mapping to Inflammatory Response. The transcripts with the most direct and indirect connections in this network was IL1β a potent pro-inflammatory modulator. IL1β and many genes in this network were only slightly above the threshold for differential expression between Ecran and Makona (>±1log2FC), highlighting the close similarities in the host response to these closely related viruses. IL1β transcript abundance was higher in Ecran than Makona infection at 24h (1.3 log2FC) and 48h (1.68 log2FC), but not significantly different by 72h (-0.4 log2FC), where both show >3 log2FC compared with mock infection. This indicated that the induction of IL1β in Makona infection was either reduced or delayed compared with Ecran at earlier time points.

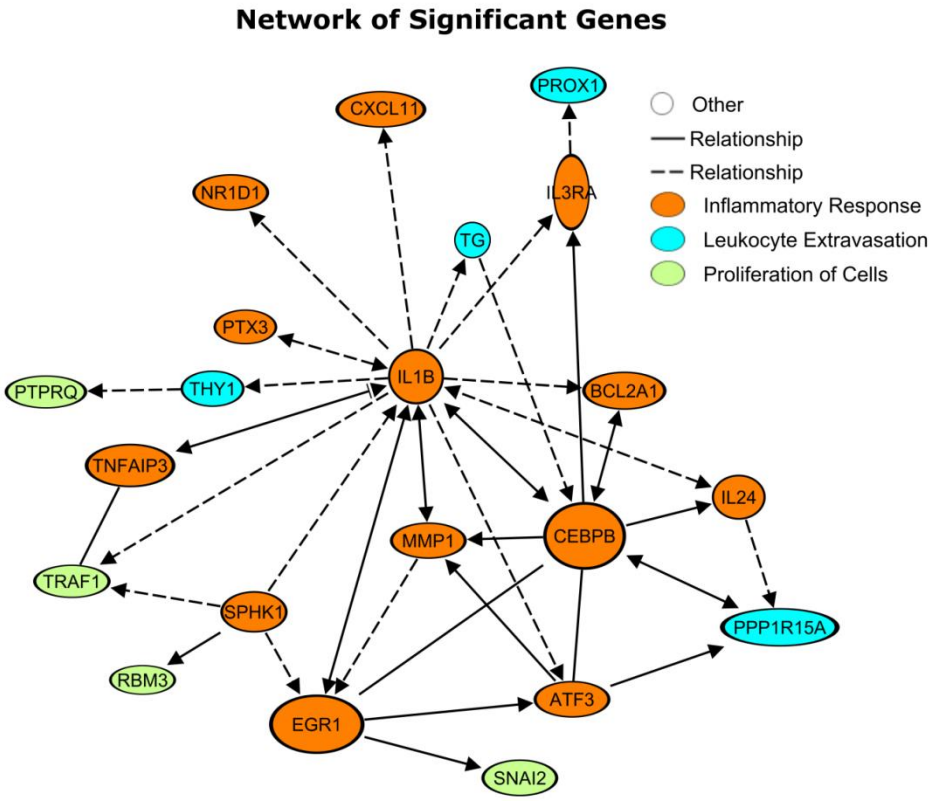


Figure 3.18: Network of significant genes identified using IPA (Qiagen). <1% of the total genes identified appeared significantly changed comparing Makona with Ecran. These genes were analysed for known associations and interactions using ingenuity pathway analysis. Genes with strong associations and experimentally verified interactions were selected for further scrutiny and a sub-network of genes with direct and indirect interactions was generated. This network is shown above. The genes shown were those which appeared different in Makona infection of A549 cells compared with Ecran. Solid line arrows indicate a direct interaction; dotted line arrows indicate an indirect interaction via intermediary proteins. Genes are shown in coloured rings. Colours correspond to functional categories of Inflammatory Response (Orange), Leukocyte Extravasation (Teal) or Proliferation of Cells (Green) based on hierarchal analysis of gene ontology.

3.6 Conclusions

The aims of the work described in this chapter, were to compare the host response to the Makona and Ecran variants of EBOV using transcriptomics in a model human cell line. Additionally, a case of EVD imported into the UK was analysed using transcriptomics. The host response profiled in the UK bound case of EVD illustrated the close relationship of viral load with the intensity of the interferon response. This also illustrated that detectable activity in the interferon response does not mean a clinical case will have a better or worse outcome, as this case was understood to have a severe EVD clinical course. Interferon signalling activity was increased during infection, as has been observed in previous studies of clinical EVD (129).

A detailed transcriptomic analysis was performed, profiling the host response to the recently emerged Makona variant of EBOV, comparing the response to that induced by the Ecran variant of EBOV. A549 cells were selected for this experiment to their known utility for studying antiviral response. This analysis showed that there were very few detectable differences in the host transcriptome during infection with these viruses. The similarity in genome sequence was similar to that previously described in a wide range of studies. Additionally, the interferon response was similar, indicating these viruses were no more or less capable of antagonising the interferon signalling pathway.

Therefore the Makona and Ecran variants should be considered very similar viruses. To further assess whether Makona induced cellular responses different from other *ebolaviruses* of different pathogenicity, a similar experiment was performed, contrasting the induced host response to Makona with that induced by RESTV to investigate the antiviral response and identify unique signatures of infection. This work is discussed in detail in Chapter 4.

4. Contrasting *in vitro* infections with the Ebola virus, Makona variant with Reston virus.

4.1 Introduction

4.1.1 Reston virus: the *ebolavirus* non-pathogenic in humans

Outbreaks of EVD caused by RESTV have been recorded in animal laboratory settings, non-human primate facilities and most recently; pig farms in South-East Asia (26). The ability of RESTV to apparently infect humans, pigs, bats and non-human primates suggested RESTV has a wide host range, but host preference is as yet undetermined. The assumption that RESTV is entirely non-pathogenic in humans does not negate the risk of epizootics in domesticated pigs (26, 29) or spill over events from suspected bat transmission (37, 38); therefore until conclusively shown to be non-pathogenic in humans a presumptive risk of zoonotic transmission should remain, and the virus handled in ACDP Containment Level 4 laboratories. This potential risk should be addressed through further investigation of RESTV. Studies of RESTV and EBOV infection in non-human primates demonstrated patterns of tissue pathology which were similar (28, 132, 191, 207), raising the possibility that RESTV has similar immune-pathologic potential.

4.1.2 Interferon response

Previous studies have compared EBOV and Marburg virus (MARV) to RESTV in cell culture to profile differential transcription of genes (136).

In vitro comparison of this kind has focused primarily on the immune response, particularly interferon regulation. These studies have demonstrated that RESTV VP24 and VP35 share interferon antagonism with EBOV albeit with reduced effectiveness (63, 101, 208).

4.1.3 Contrasting the host response to EBOV and RESTV

Comparison of *ebolavirus*es allows identification of unique signatures of viral infection, which in turn can inform design or repurposing of treatments. The recently isolated Makona variant of EBOV was suspected of displaying similar disease and replication rates, but differing pathogenesis and virulence compared with previously isolated variants (191). RESTV has previously proven to be a useful model for comparative analysis. RESTV has never caused disease in humans despite evidence of exposure (28).

4.1.4 Chapter aims

In Chapter 3 the host response to the Makona variant was considered similar to that induced by other variants, based on experiments comparing the Makona and Yambuku-Ecran variants of EBOV in cell culture. To better highlight host response changes which may be pathogenically relevant, a study was carried out to compare the host response of the Makona variant to RESTV and determine whether infection with RESTV and the Makona variant of EBOV elicited similar transcriptional patterns of antiviral and immune response genes. This was done by contrasting the replication of RESTV and EBOV Makona in multiple cell lines and comparing the cellular host response using high resolution transcriptomic analysis.

4.2 Chapter Methods

4.2.1 Infection of cell lines derived from multiple species

Cell lines derived from pigs (ESK-4, PK-15), bats (TB1-lu), non-human primates (Vero) and human (A549) were infected with RESTV and EBOV at an MOI of 0.5. These cell lines were selected because of availability of validated and authenticated cell stocks from the European Collection of Authenticated Cell Cultures, Porton Down, UK. Vero cells were selected as a non-human primate cell line, as this cell line is routinely used to culture filoviruses both in the UK and internationally for viral isolation and study. Infected cells were cultured for 72h and RNA was collected at 6h, 24h, 48h and 72h post infection. A full description of work performed at containment level 4, RNA preparation and q-RT-PCR analysis with calculation of copies/µl using standard curve is available in Chapter 2. Still cell line images were collected at x40 optic using an EVOS microscope (Thermo Scientific).

4.2.2 Viral sequencing and analysis

RNA for sequencing analysis was prepared and sequenced on an Illumina HISeq 2500 as described in Chapter 2. The detailed bioinformatics analysis method is described in full in Chapter 2. Illumina reads mapping to viral genome were analysed using QuasiBAM, a Public Health England tool to measure viral diversity in a read population.

This analysis showed diversity estimation and computed read depth for every nucleotide in the EBOV or RESTV genome. This read depth can be used to assess the relative abundance of viral mRNA being transcribed from a genomic locus (151). Nucleotide variations in EBOV genome were calculated with the assistance of CGR, University of Liverpool. Cellular transcript

abundance was calculated using Illumina reads mapped against version 38 of the RefSeq genome for Homo sapiens. Transcript counts were calculated using EdgeR in the R environment and analysed by principle component analysis, differential transcriptomic analysis and hierarchal and K-means clustering. Chapter 3, sub-section 3.5.3 shows dispersion of data.

4.2.3 Transcriptomic analysis reveals large numbers of A549 transcripts change in abundance when infected with RESTV.

Next-generation sequencing Illumina reads were mapped to human genome as described in Chapter 2. Transcriptomic analysis was performed using EdgeR as described previously (Chapters 2 and 3) and significance of transcript abundance changes calculated by the false-discovery rate method. This revealed that ~12000 transcripts showed differential abundance at 24h, 48h and 72h post infection with RESTV compared with mock infection of A549 cells. Significant transcripts were selected based on their log2CPM compared with their relative change in abundance (logFC), depicted in Figure 4.1.

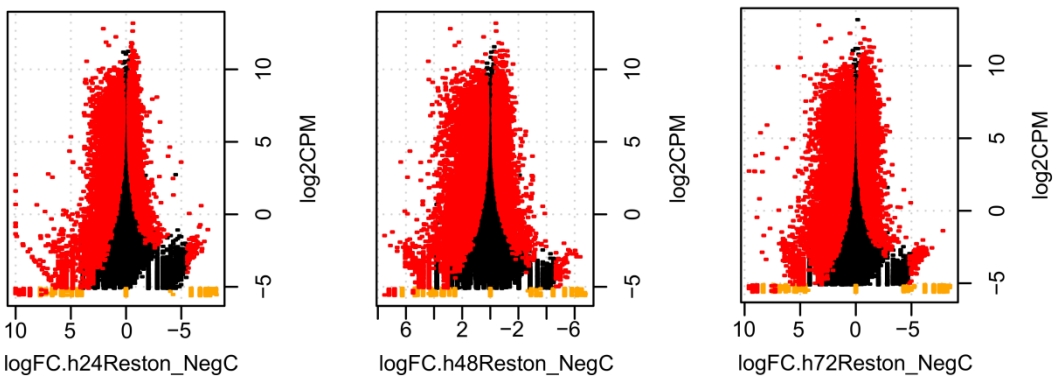


Figure 4.1: MDS plots showing the number of significant transcripts identified in this study. The plots show the logFC (fold change) in transcript abundance against the log2CPM of each transcript, which is a measurement of the average number of normalised read counts assigned to that transcript. In this way transcripts with large FC but small log2CPM may still be significant by the false discovery rate method (FDR), while those genes with small FC but very high log2CPM may also be significant. Transcripts with significant FDR are shown with in red, and those which were not significant are shown in black.

4.3 Contrasting replication of EBOV and RESTV

4.3.1 EBOV and RESTV replication was affected by cell type.

The first aim of this study was to determine whether RESTV was less able to infect and replicate cell lines susceptible to EBOV, which may help to explain the differences in pathology observed.

For comparative analysis four cell lines (A549, PK-15, TB1-lu and Vero) were infected with both EBOV and RESTV and their replication efficiency compared by validated qRT-PCR assay conducted in accordance with the method described in section 4.2.1. Multi-variate analysis using general-linear models were used and showed EBOV and RESTV genome abundance were different at all time points in A549, TB1-lu and Vero cells ($p=<0.001$). RESTV appeared less able to replicate in Vero, A549 and TB1-lu cell lines compared to EBOV; shown in Figure 4.2.

To see whether RESTV was generally unable to match EBOV in replicative efficiency, PK-15 cells were selected for further evaluation. PK-15 were the only cell lines where RESTV genome abundance was significantly higher than EBOV ($p=<0.001$) by 72h. Interestingly, RESTV appears to replicate with higher efficiency in a porcine-derived cell line, as RESTV was shown to symptomatically infect pigs in the Far East (29).

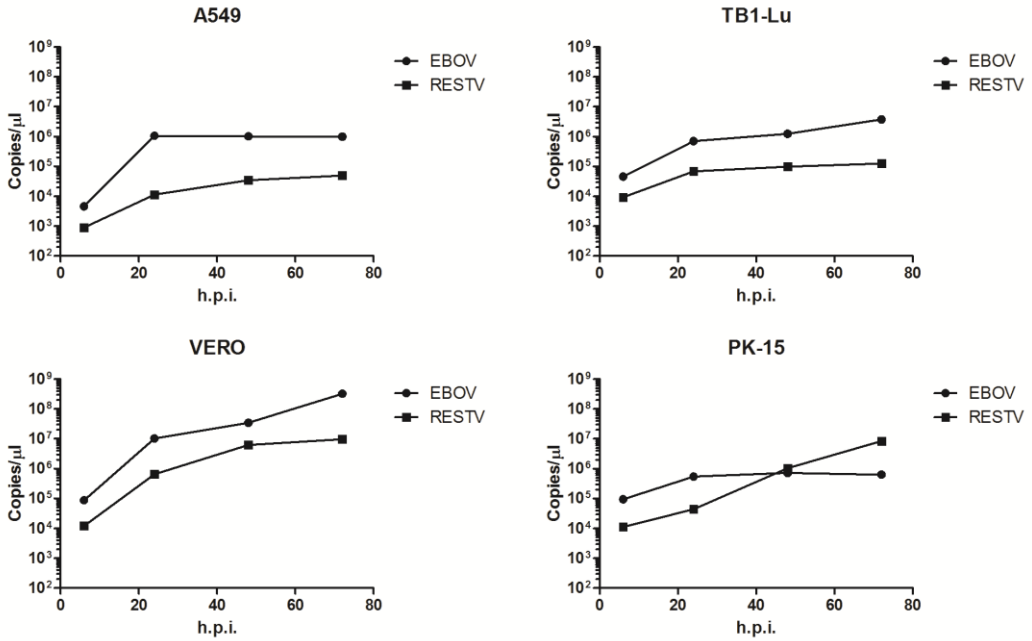


Figure 4.2: qRT-PCR analysis of viral genome abundance in cell lines derived from human (A549), bat (TB1-lu), non-human primate (Vero) and pig (PK-15). qRT-PCR values were compared with a standard curve of known copy number to calculate genome copies per microliter of extracted RNA. Data is given at 6h, 24h, 48h and 72h post infection for both EBOV and RESTV viral genomic RNA abundance. Copy number calculations were performed based on qRT-PCR data on specimens collected from 2 parallel cultures that were tested in triplicate to ensure accurate measurement (n=2). Repeated measures ANOVA was used to calculate the significance of changes in genome abundance over the time course, the Bonferoni's correction was applied.

4.3.2 EBOV and RESTV replication appeared attenuated in A549 cells

A549 cells were thought to be the strongest producers of interferon of the cells assayed, to investigate whether this may impair EBOV and RESTV replication viral RNA abundance was compared. There were significant differences in the replication of EBOV and RESTV in Vero E6, TB1-lu and PK-15 but not in A549 cells where both viruses suffered a decreased replication rate. These analyses were tested using general-linear models to assess significance with Bonferroni's correction (n=8).

4.3.3 RESTV replication was accelerated in porcine cell types.

Measurement of viral genome abundance suggested RESTV showed greater levels of abundance in PK-15 than EBOV by 72h post infection. Analysis of viral genome abundance at each time point using general-linear models showed that RESTV demonstrated greater replication efficiencies in Vero E6 and PK-15 cells, and reduced rates of replication in A549 and TB1-lu. Conversely, EBOV replication rates were greatest in Vero E6 and TB1-lu cells, and much reduced in PK-15 cells. To confirm these observations a second porcine kidney cell line was infected with EBOV and RESTV at an MOI of 0.5. CPE developed in other infected cell lines by 72h for EBOV, but did not develop for RESTV during the time course experiment. However in this case RESTV produced more severe CPE than EBOV by 72h post infection (Figure 4.3).

A further porcine cell line was selected to verify the initial hypothesis that RESTV showed accelerated replication in cell lines derived from pigs. ESK-4 cells are similar to PK-15 in that they are derived from porcine kidney cells. Viral genome abundance was measured at 6h, 24h, 48h and 72h post infection using qRT-PCR, and rates of viral replication evaluated over the time course. This analysis showed reduced levels of viral genome abundance in ESK-4 cells compared with previous cell lines tested.

However, as seen in PK-15 cells RESTV displayed greater replication in ESK-4 cells than EBOV at 24h, 48h and 72h. The significance of these differences were analysed at each time point (n=2) finding that RESTV displayed higher viral genome abundance by 72h post infection compared with EBOV (Figure 4.4). This combined analysis suggested RESTV was better able to cause productive infection of porcine cell lines compared with those derived from other species. Conversely this analysis also suggested EBOV was better able to replicate in bat cells than other cell types assayed, and that A549 cells appear to be least supportive of EBOV and RESTV replication.

4.3.4 Porcine cells show greater increase in interferon associated transcript abundance in EBOV infection compared with RESTV infection.

Measurement of viral genome abundance by qRT-PCR demonstrated that RESTV genome abundance was greater than EBOV genome abundance in ESK-4 cells. Differences in the interferon response may be the cause of this observation. qRT-PCR was used to measure the relative abundance of transcripts encoding antiviral genes. The abundance of transcripts associated with cell activation and NFκB signalling, those associated with the interferon response, and those encoding interleukins were compared. Transcripts encoding cell genes involved in the activation of the immune response and inflammation showed consistently greater transcript abundance in RESTV infected ESK-4. Transcript encoding IKBKB and NFKBIA were similar in both EBOV and RESTV infected ESK-4. Two major components of the NFκB complex, RELA transcripts (encoding p65) showed a higher abundance in RESTV infected cells and NFKB1 (p50) showed similar levels in both EBOV and RESTV.

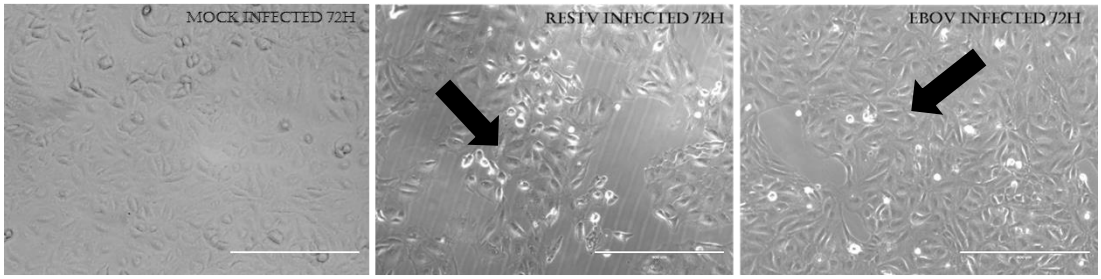


Figure 4.3: Light microscopy images (x40 optic Life Technologies EVOS microscope) of ESK-4 porcine derived cell line infected with EBOV or RESTV for 72h and an example Mock infected culture after a 72h culture period.. EBOV induces extensive CPE by 72h post infection of other cell lines but appear reduced in ESK-4 cells. Conversely, RESTV infected cells typically display no evidence of CPE until day 6, however in ESK-4 cells CPE was extensive. CPE is indicated by the rounding of cells (shown as rounded cells with a bright white halo in the above images, indicated by an arrow).

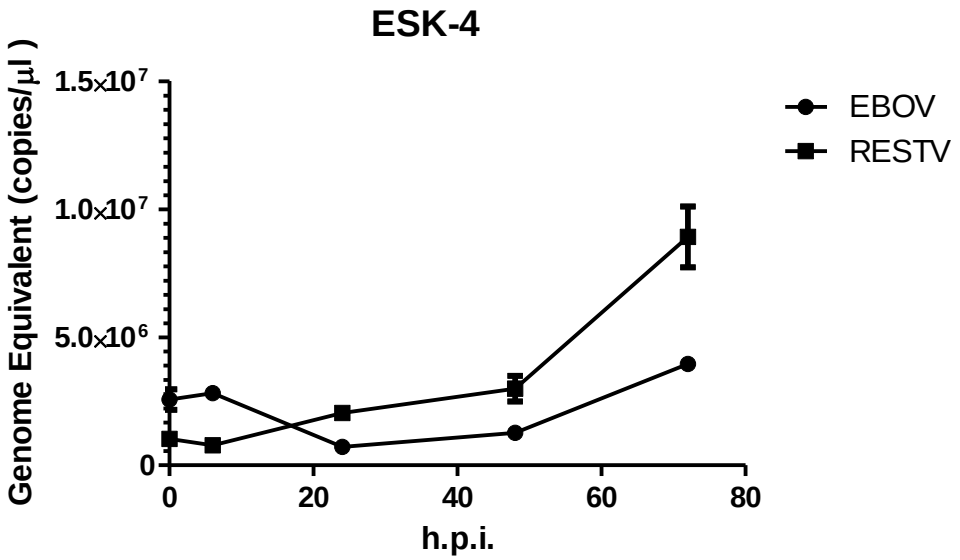


Figure 4.4: qRT-PCR analysis of a tissue culture assay of ESK-4 (porcine derived kidney) cell line infected with either EBOV or RESTV for 72h at an MOI of 0.5. RNA was collected for analysis at 0h, 6h, 24h, 48h and 72h. Viral genome abundance was calculated by measuring qRT-PCR cT values and comparing these to a standard of synthetic transcript RNA with defined copy number. The difference between 0h and 24h were minimal, while increases were seen at 48h and 72h. In this analysis EBOV genome abundance appeared lower than RESTV genome abundance at 24h, 48h and 72h.

The abundance of transcripts encoding IFNB1 and IFNAR1 were greater in EBOV than in RESTV. Lower IFNB1 transcript abundance may suggest a reduced activity of transcriptional activators associated with IFNB1 production in RESTV infected cells.

To assess whether interferon signalling was otherwise affected, transcript abundance of STAT1, IRF3 and IRF7 were analysed, showing that the abundance of these transcripts were greater in RESTV than in EBOV. Additionally, other ISGs (MX1, ISG15 and OAS2) all showed similar transcript abundance in RESTV and in EBOV infection. The production of inflammatory factors was also assessed, to determine whether RESTV invoked a similar inflammatory response in ESK-4 cells compared with EBOV. IL8 and IL1B1 transcripts were of similar abundance in RESTV infection compared with EBOV, though transcripts encoding the potent pro-inflammatory interleukin IL6 was more abundant in RESTV infected ESK-4 cells, and the abundance of transcripts encoding the pro-inflammatory regulator RELA were also notably higher in RESTV infection compared with EBOV. The results of these analyses were shown as graphs in Figure 4.5.

This analysis suggested that interferon signalling could not clearly explain the ability of RESTV to replicate better than EBOV. The role of NFκB signalling and cellular activation appeared to impact on virus replication. Furthermore, the strength of NFκB activation may correlate with the ability of EBOV and RESTV to replicate in susceptible cell lines, and this is explored later in this chapter, and in later chapters (Chapters 5 and 6).

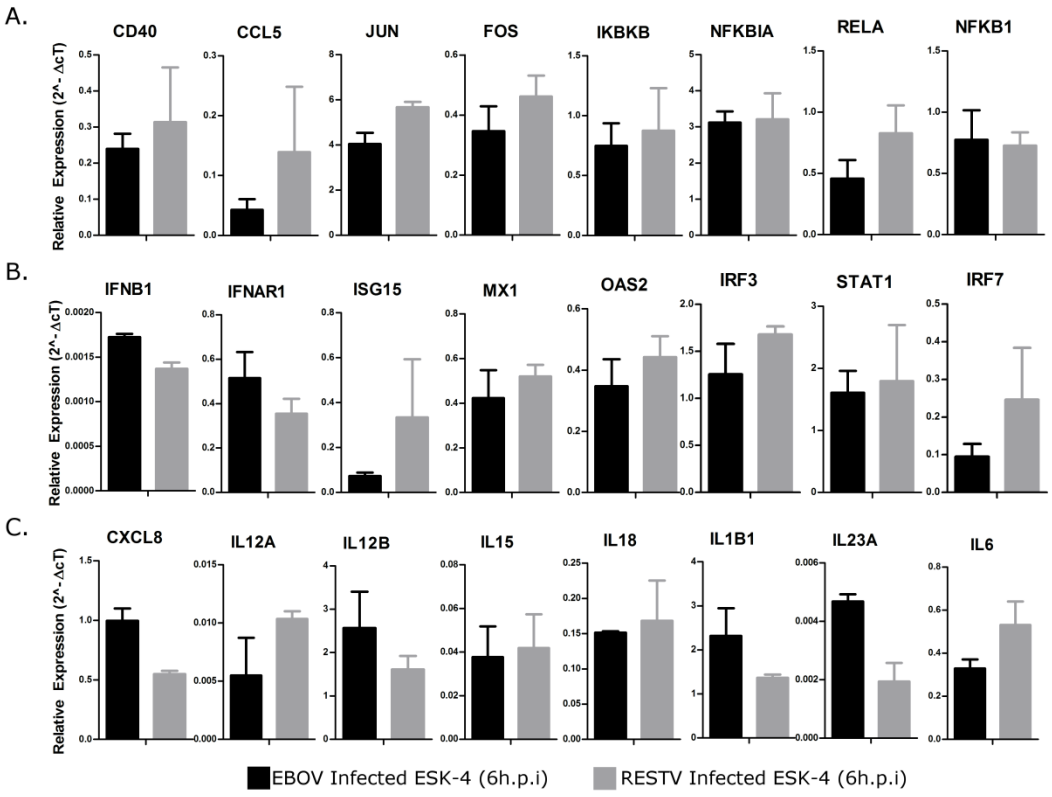


Figure 4.5: Measurement of antiviral transcript abundance changes in ESK-4 cells infected with EBOV or RESTV, 6h post infection. qRT-PCR was used to assess the transcript abundance of three subsets of antiviral transcripts. (A) The abundance of transcripts associated with the inflammatory response and activation of the immune response are shown, where most genes displayed show greater abundance in RESTV infection than EBOV. (B) Interferon stimulated genes are displayed, there appeared to be greater abundance in most genes in this subset, though IFNB1 transcript abundance was greater in EBOV than in RESTV (in contrast to A549 cells discussed in chapter 3). (C) Abundance of transcripts encoding interleukin genes. Values shown are the average delta cT values based on duplicate infections. Data from cells infected with EBOV is displayed in black and RESTV in grey.

4.4 Host response to EBOV and RESTV in A549 cells

4.4.1 EBOV and RESTV transcript abundance does not conform to expected patterns in A549 cells

An investigation into the EBOV transcriptional profile observed in infected A549 cells was shown in Chapter 3. This analysis revealed that the mapping pattern of sequencing reads did not indicate a trend in viral genome abundance which correlated with an expected gradient pattern of transcription.

To see whether this phenomenon was also detected in RESTV infected A549 cells the read depth across the EBOV, Makona and RESTV genomes were contrasted; revealing that similar to that reported in Chapter 3, a spike in read depth was visible at the position of the viral GP. This analysis also shows that the expected gradient reported in previous studies of polymerase activity in Mononegavirales does not appear to be present. The results of this analysis are shown in Figure 4.6. This discernibly greater proportion of GP transcript present in the analysis may be due to a true spike in transcriptional activity, or more plausibly may be due to the GP transcript displaying much greater stability than other ebolavirus gene products, and this should be investigated further as it is an as yet undescribed aspect of ebolavirus biology, and the contribution this would make to lifecycle and pathogenesis are unknown.

In this analysis the genomic region encoding the L gene of EBOV and RESTV appeared to show limited read depth, while the GP genomic region shows the greatest read depth. This analysis did not account for the presence of both genomic and complementary transcripts, but the diverse array of read depth could be explained by differing transcriptional profiles across the genome.

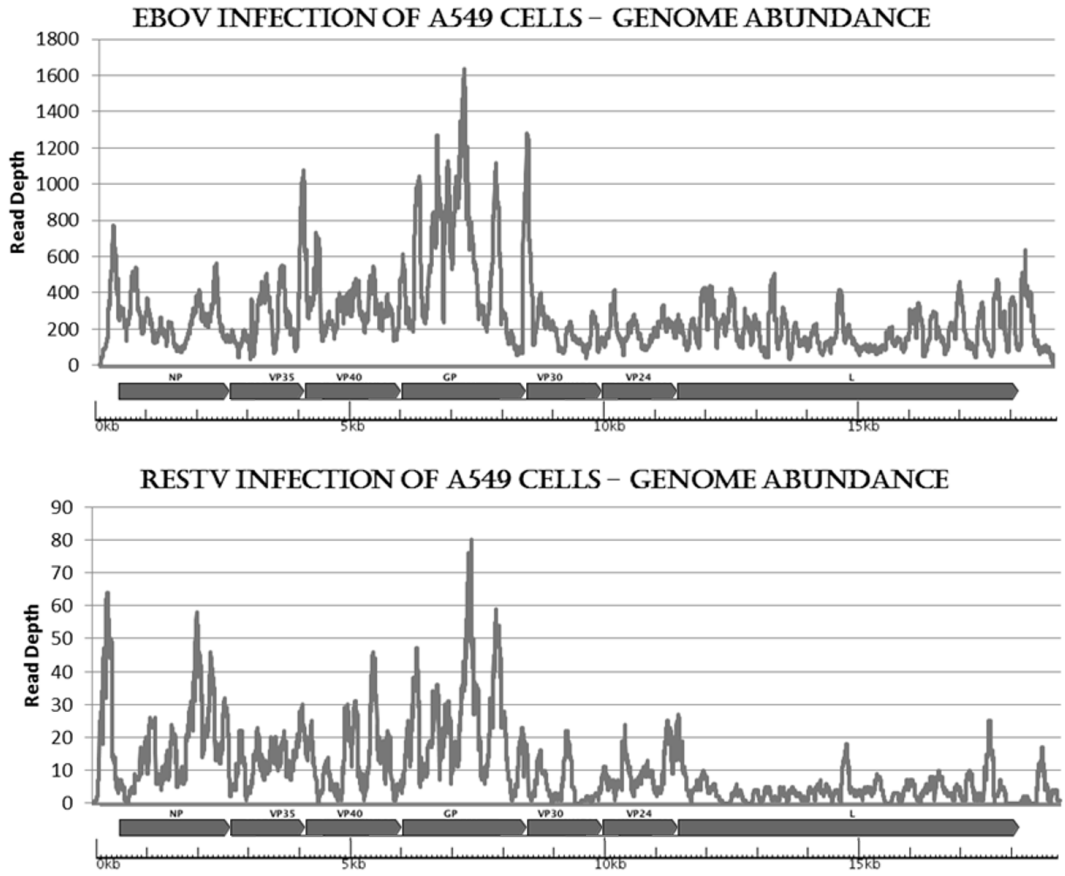


Figure 4.6: Viral RNA genome and mRNA transcript abundance was calculated by analysing the levels of read depth at each nucleotide of the viral genome. This analysis indicated that the most depth was detected over the GP (glycoprotein) gene of both EBOV and RESTV at 72h in A549 cells. The transcriptional gradient expected in Mononegavirales was not present. An abundance change was identifiable at the junction of the VP35/VP40 and the VP30/VP24 viral gene in RESTV, where gene boundaries were different from EBOV. The L gene shows the least mapping. The y-axis shows the read depth at each nucleotide and the x-axis displays the nucleotide position and the architecture of the viral genome for easy interpretation.

4.4.2 The cellular response to EBOV and RESTV were most different at 24h post infection.

Contrasting EBOV variants Makona with Ecran showed that most differences in the host transcriptome were identified at 24h post infection (Chapter 3), suggesting differences in transcriptomic profile were clearest at the earliest sampling time.

A principle component analysis was performed, illustrating that the most divergent transcriptome datasets comparing EBOV, Makona with RESTV were sampled at 24h and 48h post infection. Figure 4.7 illustrates that mock infected A549 cells can be easily separated from EBOV and RESTV infected A549 cells by the 2nd and 3rd component, or the 1st and 2nd component.

The 24h and 48h post infection host response data appeared most similar at 72h post infection when separated by all components. The 24h time point data was markedly more divergent that the 48h time point data comparing EBOV with RESTV.

For this reason the 24h time point became the focus of study, to improve the chance of identifying antiviral pathways and genes which displayed differing transcript abundance.

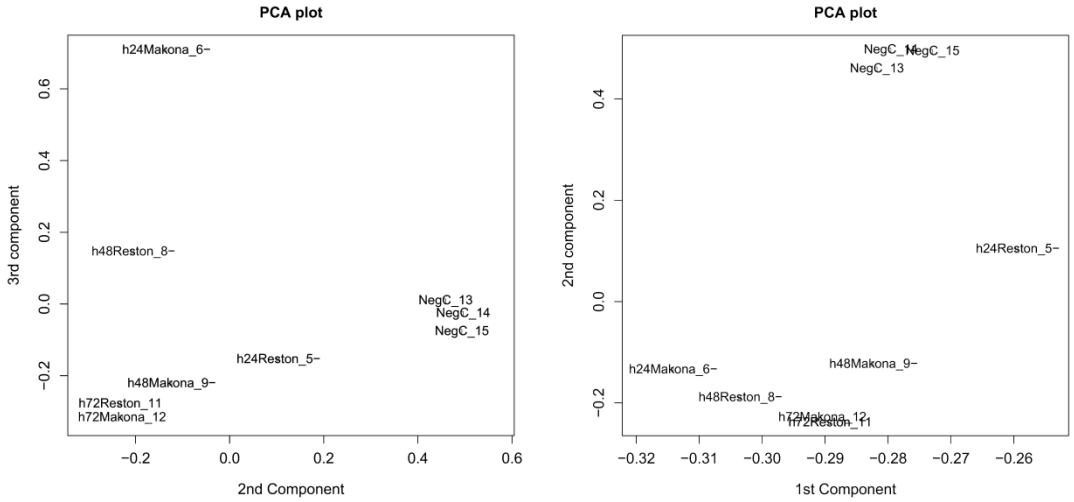


Figure 4.7: Principle component analysis of data collected from EBOV and RESTV infections of A549 cells at 24h, 48h and 72h post infection. PCA included separation by the 1st and 2nd component, and by the 2nd and 3rd component. This analysis shows that all components allow separation of data from mock infected cells from data derived from infected tissue cultures. This data also indicated that data from EBOV and RESTV infection correlates poorly at 24h, but was better correlated at 48h and very similar at 72h post infection.

4.4.3 Pathway analysis shows that EBOV and RESTV elicit similar effects on canonical pathways.

To better understand the biological significance of transcript abundance changes, canonical pathway analysis was performed. The aim of this was to identify pathways which showed significant increases in predicted activation state (z-score) elicited by infection with RESTV or with EBOV. IL-6 signalling, Acute phase response signalling, Activation by IRF by PRR, Inflammasome signalling and dendritic cell maturation displayed the highest activation states in EBOV and RESTV infection. Pathways involved in the inflammatory response were investigated: NFκB signalling was activated by both EBOV and RESTV (z-score 3.47 and 3.16 respectively). Chemokine signalling also increased in activity in EBOV (z-score 2.5) and RESTV (z-score 1.7). Pathways governing coagulation were investigated: prothrombin activation time, a pathway containing genes important in early coagulation, showed reduced levels of activity in EBOV (z-score -2.2) and RESTV (z-score -2.2). A reduction in the activity of cell cycle associated pathways (Cyclins and Cell Cycle Regulation; z-score -2.5), apoptosis (z-score -1.4) and interferon signalling (z-score -1) was found in EBOV infected A549 cells. Pathways under the regulatory control of interleukins were shown to be markedly increased in activity in EBOV and RESTV infection; these were IL-8 signalling (z-score 4.1 and 3.1 respectively), IL-1 signalling (3.2 and 2.7 respectively) and IL-6 signalling (z-score 1.1 in both infections).

4.4.4 Analysis of upstream regulators suggested that key regulators of response genes may be more active in EBOV than RESTV infection.

Pathway analysis suggested that inflammatory response pathways showed similar activation in EBOV and RESTV infected A549 cells. To confirm this observation the activation state of upstream regulators were calculated.

A negative z-score indicated greater activity in EBOV infection of A549 cells than in RESTV. FAS was identified as a significant upstream regulator with a predicted z-score of 2.213, showing greater activity in RESTV than in EBOV. Interleukin signalling was also shown to be significantly increased. The activity of IL1β was increased by a z-score of 0.444 in RESTV compared with EBOV. Pathways involved in inflammation were also identified. TNF showed lower activity in RESTV than in EBOV (z-score -0.6). SP1, and the AP-1 component JUN also showed increased activity in EBOV infection compared with RESTV (-2.4 and -2.2), NFKB1/NFKB1A and IKBKB, involved in activation of NFκB signalling were all more active in EBOV infection than in RESTV (-0.8 / -1.7 respectively).

4.4.5 The abundance of transcripts encoding immune response genes were similar in EBOV and RESTV infected A549 cells

Pathways involving the activity interleukins, chemokines, interferon, NFκB and TNFα were identified in canonical pathway analysis. To better understand they patterns of transcript abundance in these categories, transcripts downstream of key antiviral regulatory genes (IFNA, IFNB) were identified and their abundance profiles analysed, shown in Figure 4.8.

This analysis demonstrated that most interferon transcripts selected showed an increase in abundance during EBOV and RESTV infection compared with mock infection (particularly ADARB1 and ADARB2 transcripts, IFITM10, IFI44 and OAS2). The EBOV restriction factor tetherin (BST2) shows a

marked reduction in transcript abundance at 24h post infection with RESTV, but little change in EBOV (RESTV log2FC -4). Overall key interferon stimulated genes displayed similar transcript abundance profiles in EBOV and RESTV infected A549 cells.

4.4.6 Pro-inflammatory response transcripts show increased abundance in both EBOV and RESTV infection

The regulators NFκB, LPS:TL4 and TNFα were predicted to display slightly higher activity in EBOV infection than in RESTV. Additionally, transcripts involved in associated pathways showed slight differences during infection with the Makona variant in EBOV compared with the Ecran variant in A549 cells (Chapter 3). NFκB associated transcripts also appeared increased in abundance during the most severe stages of clinical EVD (Chapter 3). Therefore determine whether this translated into differences in transcript abundance, important mediators of the NFκB and TNFα pro-inflammatory response were investigated, (Figure 4.8).

Chemokines identified in the analysis showed similar patterns of expression throughout the time course in both EBOV and RESTV infection though the MIP-3α receptor CCR6 displayed significant increase in abundance (>2 FC) at all time points.

Seven NFκB associated transcripts were assessed for transcript abundance changes. Most of these complex components do not change significantly, however transcripts encoding Interleukins and NFκB components REL and NFKB1Z showed significant increases in abundance across the time course in both EBOV and RESTV infection. NFKB1Z mediates the cellular response to TLR4 stimulation, and is a promoter of IL6 transcription. Potent pro-inflammatory interleukins (IL15 and IL6) showed a sustained increase in abundance at all three time points in both EBOV and RESTV, though these

were greater in EBOV than RESTV. IL6 has been previous associated with severe EVD in human cases (124) and these results suggest an NFKB1Z mediated increase in IL6 transcript production.

In contrast RELB appeared to gradually decrease in transcript abundance over the time course exceeding -1 log2FC by 48h in both EBOV and RESTV infection. Transcripts encoding the pro-inflammatory IL1B were less abundant in RESTV infection at 24h than EBOV infection (EBOV -0.48, RESTV -4.03), and were highly increased by 72h post infection (EBOV 3.6 and RESTV 2.5).

Transcripts encoding pro-apoptotic factors show little change in abundance, but anti-apoptotic factor BCL2 show an increase in abundance in transcript abundance. Some pro-apoptotic TNFα related genes displayed increased transcript abundance. TNFSF15 is an apoptotic cytokine affecting endothelial cells showing around +2 log2FC in both EBOV and RESTV at 72h, though this was unchanged earlier in the time course. Curiously TNFAIP6 (TSG6) shows a marked reduction in transcript abundance (<-4 log2FC) in both EBOV and RESTV infection. Though this was sustained at both 48h and 72h post infection with EBOV, while in RESTV infection this only manifests at 72h. TSG6 interacts with components of Inter-α-trypsin inhibitor (IαI) to mediate activation; IαI is an important protein in the protease cascade in inflammation. Genes involved in this complex showed reduced abundance at 24h in both EBOV and RESTV infection (log2FC of ITIH2 -2.2/-1.8, AMBP -0.3/-0.5 and SPINT2 -1.3/-1.2 respectively) Overall TNFα and NFκB associated transcripts appeared increased in abundance in both EBOV and RESTV infected A549 cells.

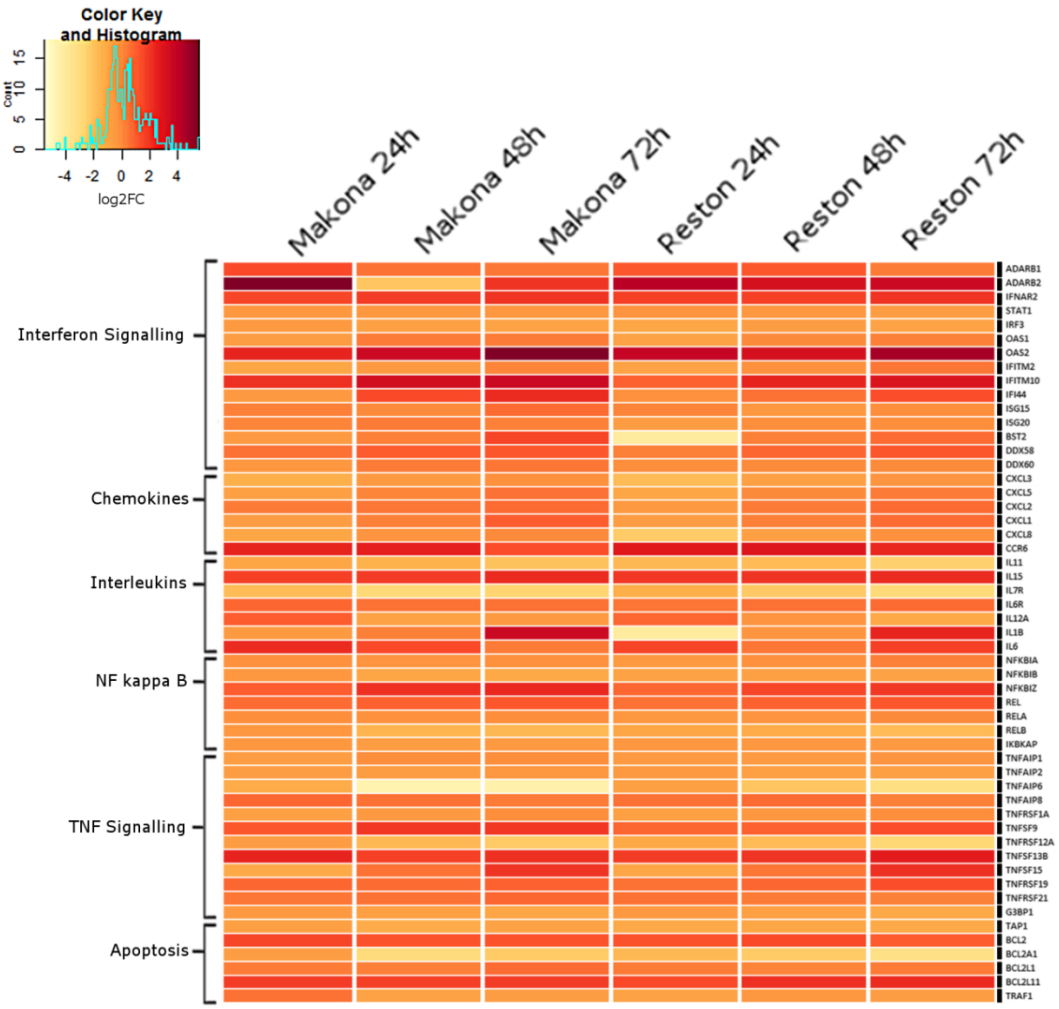


Figure 4.8: Heatmap showing transcript abundance changes across 24h, 48h and 72h post infection of transcripts in 6 functional categories. Those shown encode genes involved in interferon, chemokines, interleukins, NFκB, TNFα or apoptotic signalling. A colour scale bar is provided to aid interpretation showing change in colour from -6 to +6 with threshold values set at -4 ("Pale Yellow") and +4 ("Blood Red") for minimal and maximal colours. 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale.

4.4.7 Type I and II interferon associated transcripts were of similar abundance in EBOV and RESTV infected A549 cells

Differences in viral replication presented earlier in this chapter were considered, alongside the revelation that interferon associated genes appeared raised in both EBOV and RESTV infection. The interferon response may have resulted in differences in viral replication observed in A549 cells. Transcripts with significant changes in abundance were analysed using the Interferome v2.0 database to identify transcripts with known or predicted association with interferon response. Transcripts identified in this analysis were then annotated with transcriptome data, and displayed as a heatmap (Figure 4.9). 177 transcripts were identified in this analysis, >50% increase in abundance in both EBOV and RESTV infection, Most transcripts displayed similar abundance profiles at all time points and were type I and II associated. 12 interferome transcripts displayed significantly different abundance levels at 24h (MECOM, TYMP, CDK14, ALPK3, CSF1, CXCL2, FN1, GREM1, HIST1H3I, HNRNPDL, LTBP1, PRSS23). All of these have been associated with the interferon response or the pro-inflammatory response. Of particular note were CXCL2, a macrophage pro-inflammatory factor, FN1, a coagulation associated transcript and MECOM, associated with viral infection. Overall the analysis confirmed that the interferon response is highly active in both EBOV and RESTV infected A549 cells.

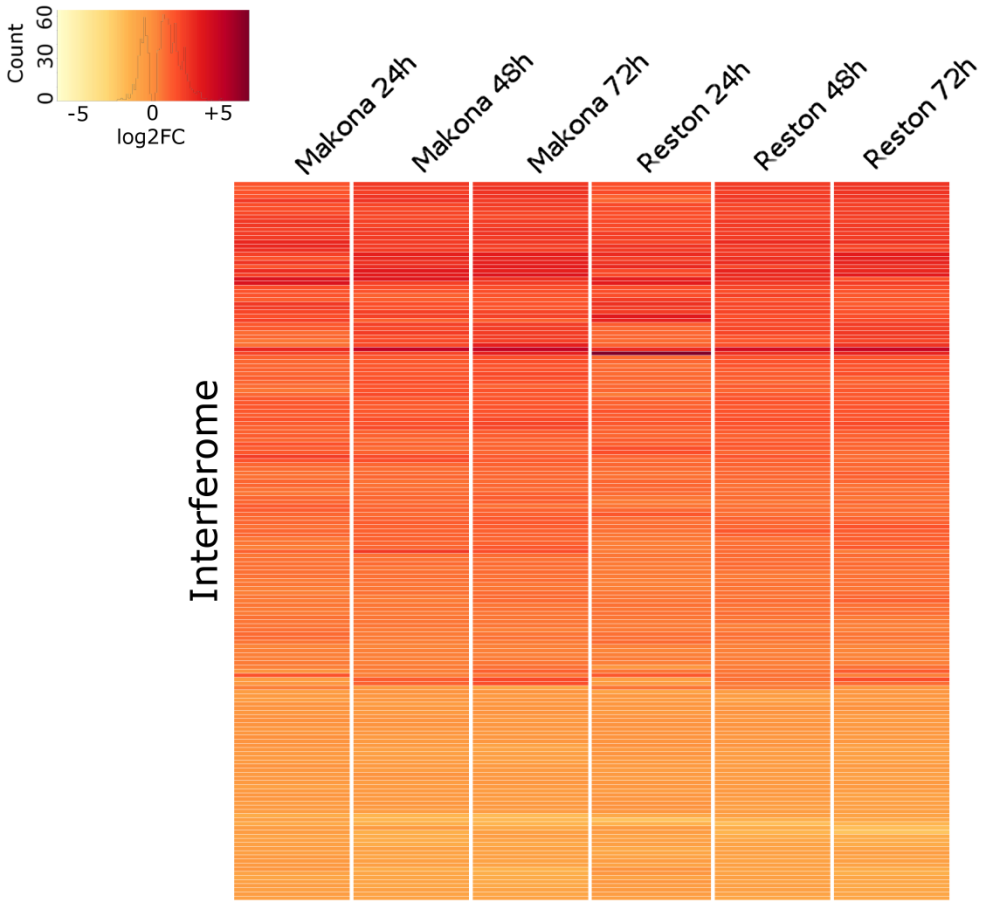


Figure 4.9: Heatmap showing the transcript abundance changes for 177 transcripts identified in an analysis using interferome v2.0. This analysis was used to identify significantly changing transcripts with a database of transcripts with known or predicted association with interferon signalling. All transcripts discovered were somehow involved in type I or II interferon signalling. >50% of transcripts increase significantly in abundance in both EBOV and RESTV at 24h, 48h and 72h post infection. A colour scale bar is provided to aid interpretation showing change in colour from -6 to +6 with threshold values set at -5 ("Pale Yellow") and +5 ("Blood Red") for minimal and maximal colours. 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into an area of the colour scale.

4.5 Conclusions

The aim of this chapter, was to compare cell infections with the EBOV, Makona variant and RESTV, and use high resolution transcriptomics to identify host response differences in EBOV and RESTV infection, and form hypothesis of their pathological relevance.

Analysis of RESTV replication in cell lines derived from human, bat, non-human primate and pig revealed differences in replication efficiencies when compared with EBOV. Additionally, RESTV showed accelerated replication in porcine cell lines. This may be due to differences in antiviral response, however the interferon stimulated transcript abundance was similar in ESK-4 cells infected with EBOV and RESTV. A dichotomy was found in the levels of inflammatory response transcripts, where RESTV infected ESK-4 cells showed greater activity in the transcription of inflammatory genes.

This chapter demonstrates that patterns of antiviral, immune and inflammatory gene transcription show significantly changed activity in EBOV and RESTV infection when compared with mock infected A549 cells. Over 50% of interferome associated transcripts display increased abundance during infection with both viruses, suggesting similarity in the induction of antiviral response.

Pathway analysis highlighted minor differences in the transcriptional activity of key inflammatory regulators in EBOV infection when compared with RESTV infection in A549 cells. To model these potentially significant changes in the inflammatory response, a cell line was selected known to be a potent expresser of inflammation associated proteins and genes. These experiments were described further in Chapter 5.

5. Modelling macrophage infection with the Makona variant virus and Reston viruses

5.1 Introduction

Infection of A549 cells with the Makona and Ecran variants of EBOV suggested only minor differences in the induction of the cellular response to viral infection (Chapter 3). A further comparative study was performed using RESTV as a control, attempting to identify characteristics of infection of A549 cells unique to the Makona variant of EBOV.

In this study, the differences in the immune response to Makona were compared with RESTV; concluding that the interferon response to EBOV Makona variant is subtly different from RESTV and other variants of EBOV. Differences in the strength of the inflammatory response mediated by NFκB, TNFα and IL1β may also be significant in the lifecycle of these viruses. Host response changes in the inflammatory response unique to the Makona variant form the basis of this chapter; a cell line useful for modelling the inflammatory response was selected to further investigate these findings in a biologically relevant cell type.

Macrophages are primary targets of *ebolavirus* infection after entering a mammalian host (6, 24, 119, 209). Macrophage derived cytokines have been detected at high levels in human patients, and may correlate with disease severity (109, 119). Secretion of pro-inflammatory mediators such as IL6 and TNFα in EVD are believed to be determined by macrophage infection (109). The effects of the Makona variant on macrophages were investigated. The aim was to identify cell biology changes unique to the Makona variant of EBOV by comparative analysis with RESTV.

This chapter describes the use of high resolution transcriptomic analysis to profile the effects of the Makona variant of EBOV on gene transcription in an *in vitro* model of human macrophages, and compares these effects to those caused by RESTV. The molecular events which take place during macrophage infection were studied using differentiated THP-1 macrophage-like cells.

A detailed description of this process is described in section 5.2.1. THP-1 cells may be differentiated in macrophage-like cells displaying phenotypic characteristics of human inflammatory macrophages (179, 180, 210, 211). The high-resolution profiles of transcription in THP-1 macrophages infected with either EBOV or RESTV are shown, and the important cell signalling events which take place during infection are discussed.

5.1.1 Chapter aims

In Chapter 4, evidence was presented that the interferon response was not significantly different in RESTV or Makona infection of A549 cells, and indicated that the interferon response was unlikely to be a major determinant of successful *ebolavirus* infection of cells, based on finding high levels of interferon production *in vivo*, and the lack of observable effect on virus replication despite high levels of ISG transcription *in vitro*.

The aim of this chapter was to characterise the host response to the Makona variant of EBOV in differentiated THP-1 cells, a macrophage-like *in vitro* cell model and compare patterns of inflammatory gene transcription in EBOV, Makona infection and RESTV infection. By using inflammatory THP-1 derived macrophages, this study sought to find differences in inflammatory response transcription between these virus infections.

5.2 Transcriptomic analysis of EBOV and RESTV infection of THP-1 cells

5.2.1 Characterising differentiated THP-1 cells

THP-1 cells are derived from neoplastic monocytes, and can be differentiated into macrophage-like cells by treating with low concentrations of a PKC activator (180). THP-1 differentiation was performed in accordance with the procedure described in sub-section 2.1.3 and cells cultured in accordance with sub-section 2.1.2. After resting they were seen to express surface proteins and morphology of pro-inflammatory macrophages (178-180) and may be further differentiated into model cells resembling the M1 (pro-inflammatory) or M2 (tissue repair) macrophage phenotypes by treating with recombinant cytokines (178). Using published protocols for THP-1 differentiation resulted in adherent cells with altered morphology. Unstimulated THP-1 monocytes were non-adherent and round, while after stimulation and resting all cells were larger, fully adherent and of irregular shape, similar to that described in previous papers (178-180, 210-216). Differentiated THP-1 cells are known to respond to viral infection with an interferon response, demonstrated with Dengue virus infection (213, 214). THP-1 derived macrophages are susceptible to *ebolavirus* infection; un-stimulated THP-1 monocytes are refractory to infection (112).

5.2.2 Replication and Infectivity of EBOV and RESTV in THP-1 cells

To test whether the growth of the two viruses was potentially equivalent (and allowing a direct comparison of the host response to be made), the replication of the viruses in cell culture and the amount of virus released was assayed by measuring the abundance of the viral genome as a proxy for this. Duplicate infections were performed at an MOI of 5 for the two different viruses in differentiated THP-1 cells.

qRT-PCR was used to measure viral genome abundance at 0h 6h, 24h, 48h and 72h post-infection. Equivalent viral genome copies were calculated by comparing the abundance of the viral genome with a synthetically generated control RNA equivalent to the GP gene of EBOV (Makona variant) or RESTV. The data was generated as genome equivalents and presented as plots Figure 5.1. Statistical analysis by two-way ANOVA showed that EBOV and RESTV viral genome abundance was equivalent in total RNA purified from cells. Whereas genome abundance in RNA purified from supernatant was approximately one log lower for RESTV compared with EBOV.

Infectivity of Makona and RESTV was also compared. Differentiated THP-1 macrophages were infected at an MOI of 0.5 or 5 and fixed in 10% neutral buffered formalin. Viral GP was detected by immune-labelling. A colorimetric staining method was employed resulting in a dense red precipitate in and around cells containing viral glycoprotein, indicative of cellular infection.

Visual counting of cells using a cell counter showed that an estimated ~25% of cells were infected at an MOI of 0.5 and over ~90% of cells infected at an MOI of 5. The number of cells infected appeared similar for both EBOV and RESTV.

Together these results indicate that intracellular viral genome abundance and viral infectivity were equivalent for EBOV and RESTV, and thus differences in transcriptomic profiles cannot be explained by differences in infectivity or replication. The observation that viral genome abundance was lower in supernatant between these viruses may suggest a reduced level of viral secretion in RESTV infection compared with EBOV, however in the absence of recoverable virus and titration, this remains to be explored further.

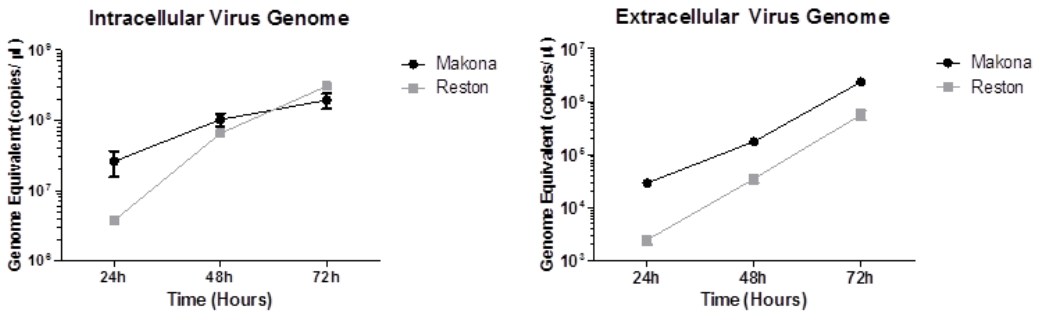


Figure 5.1: Charts of viral genome abundance detected in either Intracellular (left) or extracellular (right) extracts from virally infected THP-1 cultures. RNA was tested at 24h, 48h and 72h and genome abundance quantified by comparing cT values from qRT-PCR to a known standard of synthetically generated RNA genome fragments. EBOV (EBOV) is shown in black and RESTV is shown in grey (Reston). Values given are genome equivalent abundance values in copies of genome per µl RNA tested (copies/µl). ANOVA was performed showing no significant differences between intracellular viral genome abundance at 48h or 72h. ANOVA showed that extracellular virus was significantly different at 24h only, but visually the viral RNA levels appear lower in RESTV than in EBOV.

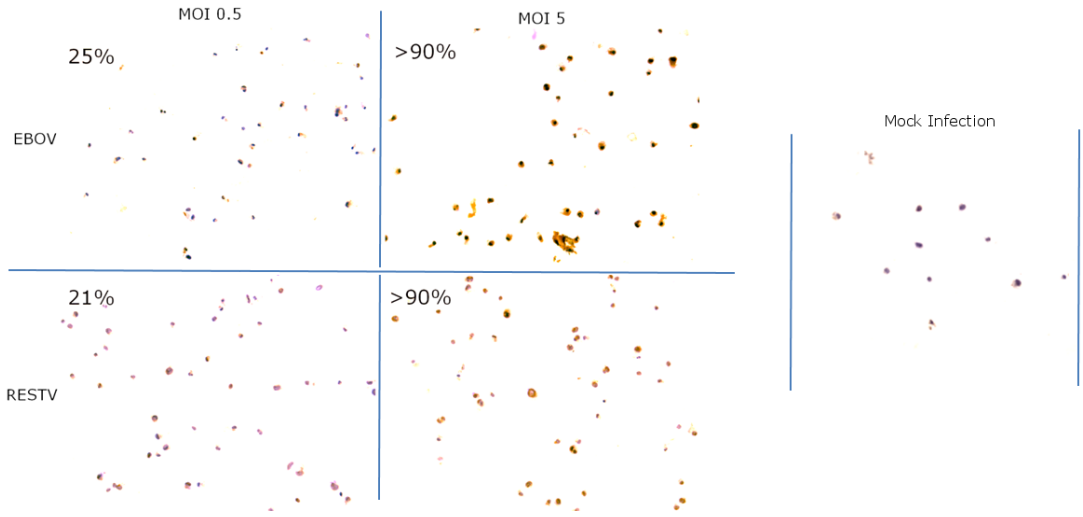


Figure 5.2: THP-1 cells were differentiated and rested before being detached from culture flasks and reseeded onto 8 well micro-well slides. Slides were infected with either EBOV or RESTV at an MOI of 0.5 or 5, or mock infected with RPMI media. After 24h post infection slides were washed and fixed in 10% NBF for 14 days. Slides were processed as described in Chapter 2: Methodologies and imaged using the automated imaging system by Histologist Laura Hunter, Public Health England, Porton Down. Cell nuclei were stained purple, cytoplasm stained pink and virus bound antibody is conjugated to an enzyme which precipitates a brown colour. Brown stained cells indicate those which were virally infected. A rough estimate of infection was performed using a cell counter and light microscopy. Slides infected with an MOI of 5 showed >90% cell infection, while those with an MOI of 0.5 showed <50% cell infection.

5.2.3 Transcriptomic analysis

To investigate host transcription in THP-1 cells infected with EBOV or RESTV, RNASeq analysis was performed with duplicate samples at 24h, 48h and 72h on total RNA prepared from the cell lysate. To establish a background level of variation in cellular transcript abundance, total RNA was prepared from 6 separate mock-infected controls and sequenced separately. Non-viral sequence reads were mapped to the human transcriptome. Data was modelled after calculating the negative binomial (NB) distribution of transcripts to better understand data dispersion. mRNA levels were normalized in order to compensate for sample size factors and general linear models (GLM) were employed (217). The TMM (trimmed mean M-values) method was used. Gene counts per million bases (CPM) were calculated using EdgeR in the R environment and FC calculated compared with mock infected THP-1 cells, or directly comparing EBOV with RESTV infected THP-1 cells. Raw reads were trimmed to remove adapter bias and the total number of trimmed, paired reads identified in each sample is shown in Table 5.3. Dispersion was estimated and applied to FDR calculations. Significant genes were identified with consideration both for CPM and the level of FC observed in samples. In this way genes with lower counts were not considered significant unless showed a large change in abundance.

This data is illustrated in smear plots shown in Figure 5.4. The number of transcripts showing significant change in abundance comparing EBOV directly with RESTV was much lower than when comparing EBOV or RESTV with mock infection.

The analysis indicated that 2085, 1815 and 2849 gene transcripts were increased in abundance and 4115, 2693 and 3902 gene transcripts were decreased in abundance at 24, 48 and 7h post infection, respectively for EBOV.

For RESTV, in total, 1912, 1849 and 1832 gene transcripts were increased in abundance and 3591, 3151 and 2793 gene transcripts were decreased in abundance at 24, 48 and 72h post infection respectively. This analysis shows similar numbers of transcripts changing overall at 24h and 48h, but different numbers at 72h. Fewer transcripts appeared to be increasing in abundance at 24h in EBOV than in RESTV. At 48h, the numbers of increasing transcripts was equivalent. At 72h the number of transcripts increasing and decreasing in abundance were both greater in EBOV infected THP-1 cells.

Sample Designation	Trimmed Read Count
Sample_1-EBOV-Makona_24h-1	97520622
Sample_2-EBOV-Makona_24h-2	93280814
Sample_3-EBOV-Makona_48h-1	90827810
Sample_5-EBOV-Makona_72h-1	91938989
Sample_6-EBOV-Makona_72h-2	96606806
Sample_7-RESTV-Sienna_24h-1	89215277
Sample_8-RESTV-Sienna_24h-2	79467688
Sample_9-RESTV-Sienna_48h-1	80447322
Sample_10-RESTV-Sienna_48h-2	92254758
Sample_11-RESTV-Sienna_72h-1	77704927
Sample_12-RESTV-Sienna_72h-2	73455770
Sample_13_Mock infected 1	74166157
Sample_14 Mock infected 2	89666547
Sample_15 Mock infected 3	66849569
Sample_16-Mock infected 4	79218617
Sample_17-Mock infected 5	65942479
Sample_18-Mock infected 6	75785700

Table 5.3: Table showing accepted paired-end reads after trimming with CutAdapt and mapping to human transcriptome hg19. Total trimmed read counts were shown alongside sample ID which includes run sample number, Virus, time point and replicate number. Read counts range from ~66 million reads to ~98 million reads.

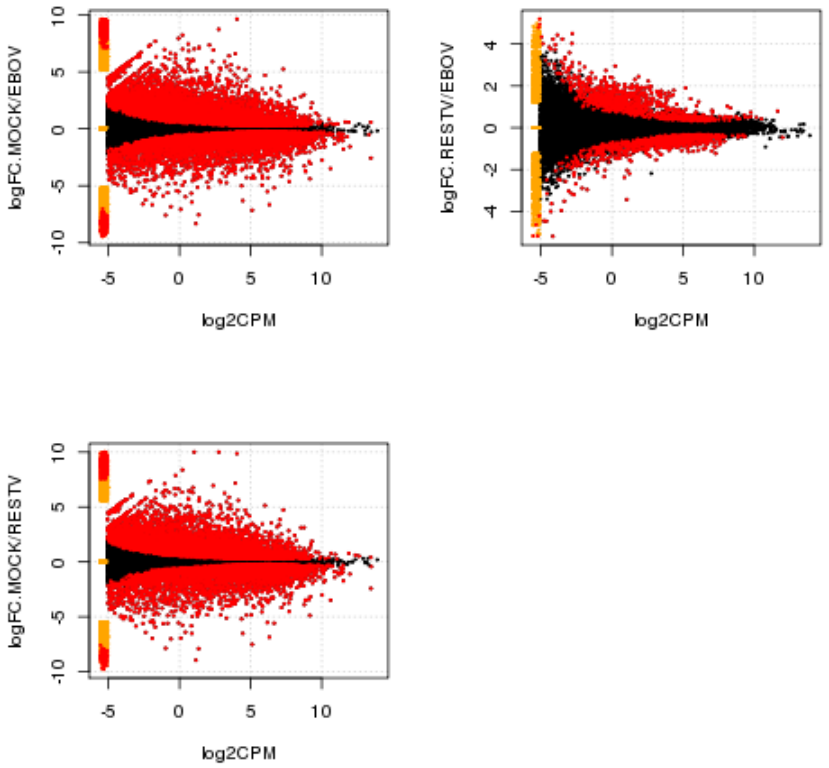


Figure 5.4: Transcript abundance was calculated and dispersion estimated in EdgeR to calculate FDR for transcripts at different log2CPM average abundance levels. This was used to identify significant genes based on average CPM accounting for FC. These plots show the number of genes identified as have significant transcript abundance changes. Red dots indicate significant transcript abundance changes, black dots signify those which were not significant, and were not included in the final analysis. FC comparing EBOV/MOCK and RESTV/MOCK were shown in the left hand smear plots. The right hand smear plot shows substantially reduced numbers of red dots (significantly changing transcripts), comparing FC of EBOV directly with RESTV.

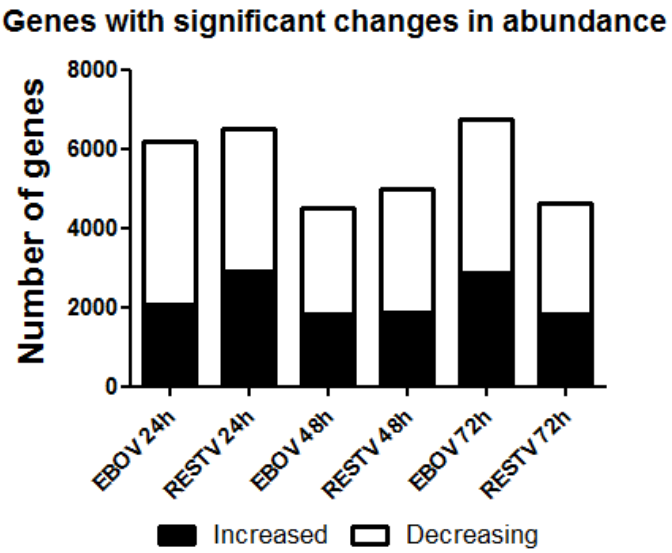


Figure 5.5: The total number of genes with significant transcript abundance changes was shown in this figure. The total values were split between 3 time points; 24h, 48h and 72h post infection. Columns show numbers of transcripts significantly increasing (Black) and decreasing (White) in abundance. This analysis shows similar numbers of transcripts changing overall at 24h and 48h, but different numbers at 72h. Fewer transcripts appeared to be increasing in abundance at 24h in EBOV than in RESTV. At 48h, the numbers of increasing transcripts was equivalent. At 72h the number of transcripts increasing and decreasing in abundance were both greater in EBOV infected THP-1 cells.

5.2.4 Viral genome and mRNA abundance and sequencing

The abundance of viral mRNA was assessed by calculating the cumulative number of trimmed illumina sequencing reads mapping to each nucleotide of the viral genome. The viral genome was assembled and mapped to the RefSeq genome for the Makona-C05 isolate or Reston-Philippines-Pennsylvania isolates using BWA (using the SPADES assembler). Mapped assemblies were then assessed using QuasiBAM, a Public Health England tool used to assess viral genome diversity. This analysis showed that diversity compared with 24h genome was <1% for both RESTV and EBOV. As described in A549 cell infections the expected gradient of viral mRNA abundance was not apparent, with most reads mapping to the GP ORF of the EBOV and RESTV genomes Figure 5.6.

5.2.5 Data modelling and statistical analysis

The RNASeq data was analysed using a number of different approaches focused on treating the data as a whole and then focusing down on the differential abundance of specific genes and potential changes in signalling pathways. Differences in gene abundance were investigated by examining the correlation between the various data sets corresponding to each condition. The data was presented as a correlation heat map Figure 5.7A and a principal component plot Figure 5.7B. The data indicated that host transcript abundance differed in infected cells from mock infected cells for both viruses.

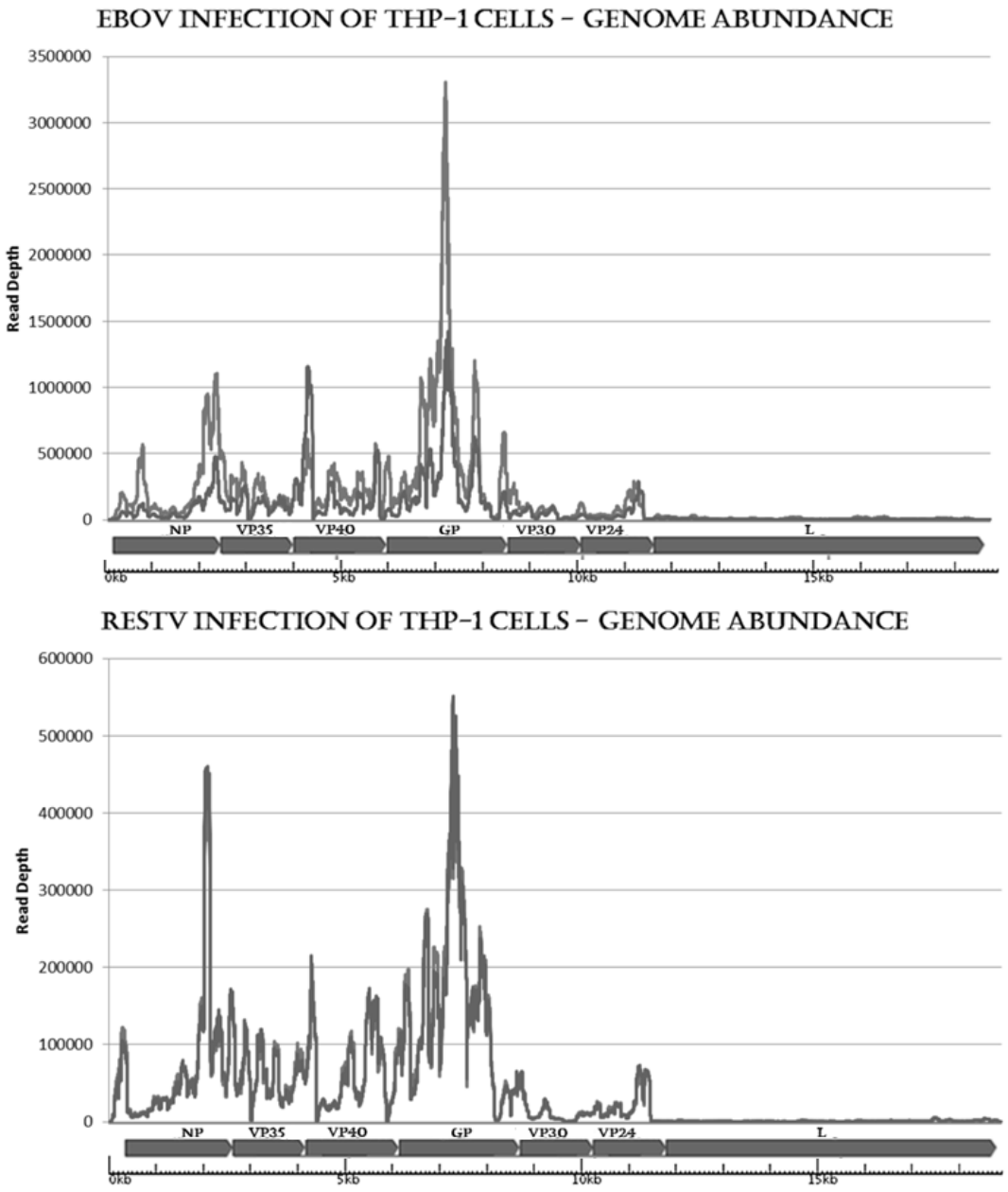


Figure 5.6: Genome coverage of illumina sequencing reads mapped to viral genome sequences for EBOV and RESTV. Displayed on the x-axis were the nucleotide positions and the architecture of the genome, showing each of the 7 ebolavirus genes sequentially from 3' to 5'. Data from two parallel experiments were displayed (Red and Blue). Y-axis shows the cumulative number of reads mapping over each nucleotide of the genome (Read Depth).

Host transcript abundance in mock infected cells clustered closely together. Thus biological variation amongst this latter group of control samples was low.

The profile of host transcript abundance for THP-1 cells infected with EBOV and RESTV infected THP-1 cells at 24h and 48h clustered closely together with low levels of variation. However, at 72h post-infection the profile of host transcript abundance for both EBOV and RESTV correlated less than at earlier time points for both EBOV and RESTV (e.g. Figure 5.7A). To investigate this further, host transcript abundance data was normalised and combined for the duplicate samples. Host transcripts with an abundance difference of 2 or more (both increased and decreased compared with the mock-infected controls) and identified with an FDR corrected p-value of <0.05 were tested for correlation Figure 5.7C.

The analysis confirmed that host transcript abundance was most divergent between EBOV and RESTV at 72 hr post-infection. Figure 5.7C shows the calculated FC comparing EBOV with RESTV infected THP-1 cells. This data indicated that the 72h time point had the most divergent host transcript abundance between EBOV with RESTV. Individually the EBOV infection 72h time point datasets correlate less than datasets from duplicates for other time points and conditions. One of the two duplicate infections at 72h was significantly different from the other. The use of negative binomial distribution to calculate dispersion of data adequately compensated for any anomalous differences in transcripts from each duplicate experiment, and those that were significantly different were not included in analysis. In contrast, earlier time points showed close correlation. Transcripts with a FC greater or lower than 2, with an FDR corrected p-value of <0.05 were used in further gene ontology and canonical pathway analysis to investigate the biological effects of these transcriptomic changes grouped into related pathways and potential functions.

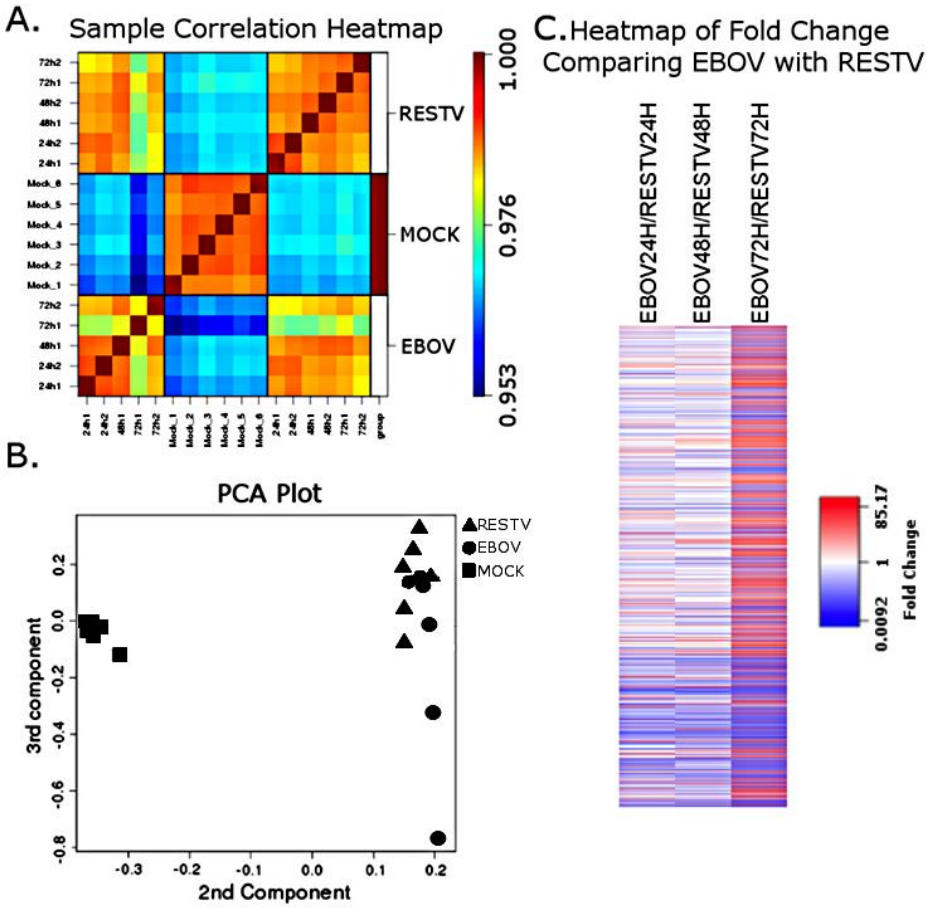


Figure 5.7: Correlation of transcriptomic data. (A) A correlation heatmap illustrating that 6 samples from mock infected cultures cluster very closely, but separately from data from infected cells. This also demonstrates that 72h time point for EBOV was more divergent than other time points for either EBOV or RESTV. A scale bar is provided from Blue (poor correlation) to Red (close correlation). (B) A principal component analysis (PCA) plot showing the statistical analysis of transcriptomic data to identify relationships in the data between each sample. Once again this analysis shows close clustering of samples from mock infected cells, distinctly from infected cells, which can be clearly separated by the 2nd component. Samples of RESTV infection better at 24h and 48h than at 72h with those from EBOV infection. Therefore the data from each sample can be separated by the 3rd component as time progresses. The 72h samples of EBOV and RESTV infections show greater separation by the 3rd component than 24h or 48h samples. (C) Heatmap showing FC comparing transcript abundance changes in EBOV directly with RESTV. A scale bar is provided with a threshold value of 0.0092 FC (Blue; logFC -6.76) and 85.17 FC (Red; log FC +6.41). This analysis shows that the most abundance changes were identified at 72h post infection, with noticeably fewer transcript abundance changes at 24h and 48h post infection.

5.3 Profiling the THP-1 host response to EBOV and RESTV

5.3.1 Interferon signalling

The abundance of selected transcripts associated with the interferon response were examined Figure 5.8. These included potential activation of STATs and IFN-stimulated genes. Comparison of EBOV and RESTV infected cells with the mock-infected cells suggested no significant increase in the abundance of STAT2 mRNA at any time during infection, though STAT1 appeared different at 72h in EBOV and RESTV infection when compared directly. MX1, OAS1, OAS2, IFIT2, IFI44 and IFIT1 showed greater transcript abundance at 72 post infection with RESTV but not EBOV. Transcripts encoding ISG15 and ISG20 also increased in abundance in both RESTV and EBOV at 72h, but not above 2 log2. IFI16 and IFITM2 transcripts all decreased in abundance in both EBOV and RESTV infected cells. Other interferon stimulated genes depicted did not show significant difference from those in mock infected cells. This suggested at the interferon response was not activated in infected THP-1 cells. To test for the expression of interferon and ensure that the THP-1 cell line used in this study was capable of expressing transcripts encoding interferon, a more focussed analysis of antiviral genes was performed, shown in section 5.4.

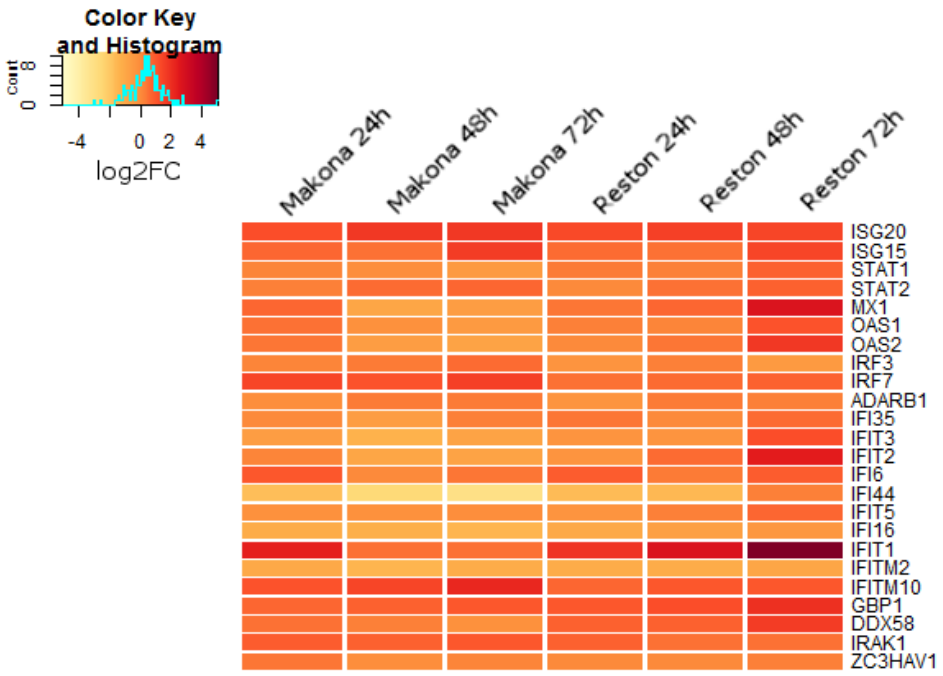


Figure 5.8: Heatmap showing log2FC in transcript abundance of transcripts encoding genes identified as Interferon Stimulated Genes (ISG) compared with mock infected cells. A colour scale bar is provided to aid interpretation showing change in colour from -5 to +5 with threshold values set at -4 ("Pale Yellow") and +4 ("Blood Red"). 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale. The histogram indicated that the greatest numbers of transcripts in this heatmap of ISG were close to 0 log2FC, with most falling in the range of -2 to +2 log2FC.

5.3.2 Interleukins and Chemokines

The abundance of transcripts encoding many chemokines were increased in response to infection with both EBOV or RESTV Figure 5.9. Chemokines were divided into 5 categories (CC, CXC, CCR, CXCR and XC). Transcripts encoding the chemokines CCL2, 3 (MIP-1α), 4, 8, 13, 19, CXCL1, 2, 3, 8 and CXCL13 showed the greatest abundance 24h after infection with either EBOV or RESTV. CXCL11 had the greatest abundance at 24h in EBOV infection but at 72h after RESTV infection. CCR1, 2 3 4 and CCL21 and 23 were all decreased in abundance in response to EBOV and RESTV infection. CCL23, otherwise called MPIF-1 is highly chemotactic for resting monocytes and lymphocytes and decreased in abundance in EBOV and RESTV infection. CCL5 also referred to as RANTES showed an increase in abundance at 24h in both EBOV and RESTV infection but remained unchanged at 48h and 72h. RANTES is an important chemotactic cytokine involved in recruitment of lymphocytes to inflammatory sites. The increased abundance of transcripts of chemokines would lead to the prediction that other pro-inflammatory cytokines would increase in abundance (e.g. IL1, IL6 and TNFα). The changes in abundance of interleukins were investigated. Transcripts encoding interleukins 6, 10, 1β, 1A, 7, 11, 15, 24, 32, 34 and 36 showed increased abundance in infected cells Figure 5.10. Transcripts encoding interleukins 12, 16, 17, 20, 37 appeared to decrease in abundance.

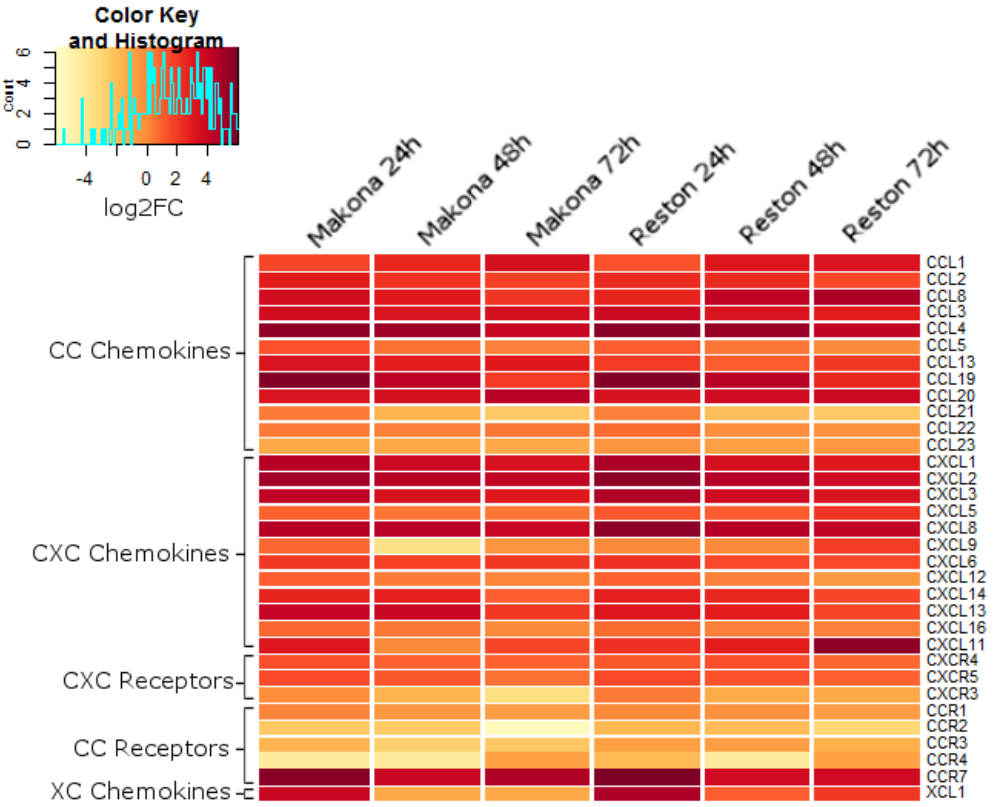


Figure 5.9: Heatmap showing log2FC in transcript abundance of transcripts encoding genes identified as "Chemokines" compared with mock infected cells. 5 categories of genes were shown, separated by chemokine classification (CC, CXC, CXCR, CCR and XC). A colour scale bar is provided to aid interpretation showing change in colour from -6 ("Pale Yellow") to +6 ("Blood Red"). 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale. The histogram indicated that the greatest numbers of transcripts in this heatmap of ISG show an increase in abundance.

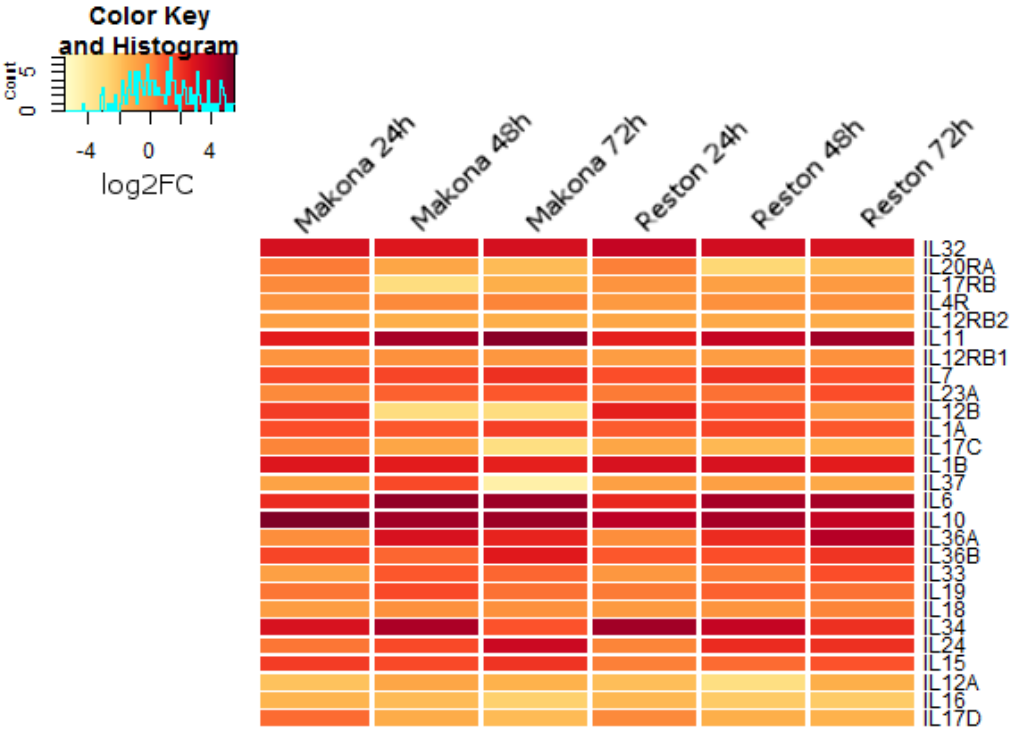


Figure 5.10: Heatmap showing log2FC in transcript abundance of transcripts encoding genes identified as "Interleukins" compared with mock infected cells. A colour scale bar is provided to aid interpretation showing change in colour from -5 to +5 with threshold values set at -4 ("Pale Yellow") and +4 ("Blood Red"). 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale. The histogram indicated that the transcript abundance changes were highly diverse and range across the scale from -4 to +4, with the most transcripts falling into the range of ~+2 log2FC.

5.3.3 TNF associated signalling

Analysis of the blood transcriptome in EBOV infected patients suggested a large acute phase reaction compared with healthy controls (218). One of the cytokines that contributes towards this response is tumour necrosis factor (TNFα). Therefore the abundance of TNFα and associated transcripts was investigated. Transcripts were categorised and divided under 5 sub-headings (TNF associated, transcription, TNF signalling, cell death signalling and signal transduction). Transcripts encoding TNFα increased in abundance 4 log2 fold for EBOV, Makona and RESTV at 24h compared with mock-infected cells, and decreased in abundance at 48 and 72h. The abundance of transcripts regulated by, or under the control of TNFα also increased. For example, transcripts produced from TNFAIP6 were increased in abundance 6 to 7 log2 fold both EBOV and RESTV at the three time points. TNF can also stimulate the expression of FOS and RELB transcripts, and these were also found to be increased in abundance in THP-1 cells infected with EBOV or RESTV. TNFSF10 (TRAIL) is a pro-apoptotic gene whose transcript was reduced in abundance in both EBOV and RESTV infected cells compared with mock-infected cells. Additionally, TRAIL transcript appeared of lower abundance in EBOV infection than in RESTV.

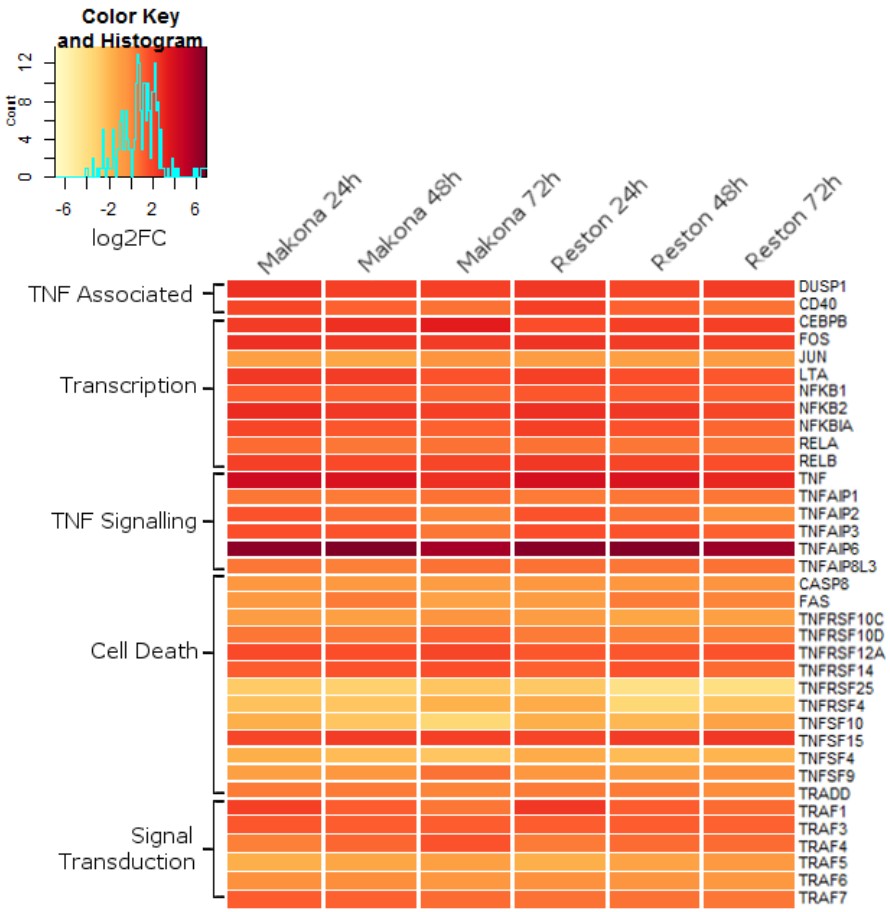


Figure 5.11: Heatmap showing log2FC in transcript abundance of transcripts encoding genes identified as "TNF Associated Genes" compared with mock infected cells. 5 categories of genes are shown, separated by functional group (TNF associated, transcription, TNF signalling, cell death and signal transduction). A colour scale bar is provided to aid interpretation showing change in colour from -6.5 log2FC ("Pale Yellow") to +6.5 log2FC ("Blood Red"). 0 log2FC indicated no change compared with mock infected cells and is indicated by "Sunset Orange". A histogram is shown in the scale bar, indicating the frequency that genes fall into a particular zone of the colour scale. The histogram indicated that the greatest numbers of transcripts in this heatmap of ISG show have log2FC values between 0 and +2 log2FC.

5.3.4 Identifying significant differences between EBOV and RESTV

EBOV in humans causes severe disease but as yet there has been no indication that RESTV causes any disease in humans. The RNASeq analysis suggested both similarities and differences in transcriptional profiles in cells infected with EBOV compared with RESTV. For example, the transcript encoding IL6 increased with similar abundance, with a log2 FC of 4.127 and 4.297 in RESTV and EBOV infected cells, compared with mock-infected cells respectively. However, differential transcriptomic analysis identified 484 transcripts with decreased abundances, and 157 transcripts with increased abundance in EBOV infected THP-1 cells compared with RESTV infected THP-1 cells, these analyses are depicted in Figure 5.12. Differentiated THP-1 cells have been used a model of human macrophages to explore whether there was a difference in the activation state of THP-1 cells, the abundance of transcripts encoding two macrophage activation and differentiation markers, arginase (ARG1) and nitric oxide synthase (NOS2), were compared. ARG1 indicated an anti-inflammatory phenotype while NOS indicated a pro-inflammatory phenotype. The abundance of transcripts encoding ARG1 was increased in EBOV infected cells but decreased in RESTV infected cells. In contrast, the abundance of transcripts encoding NOS2 was slightly decreased in both EBOV and RESTV infected cells. Transcripts encoding IFIT2, OASL and TRIM22 were increased in abundance in RESTV infected cells, but not in EBOV infected cells. Transcripts encoding the chaperone proteins, HSPA1A and HSPA1B were increased in abundance in both infections, but were greatest in RESTV infected cells. The transcripts encoding JUND, a component of the AP1 pro-inflammatory complex, were more abundant in EBOV infected cells than RESTV infected cells.

5.3.5 Pathway analysis and gene networks

Transcripts showing significant changes in abundance in EBOV infection compared with RESTV infection of THP-1 cells were analysed using Interferome v2.01. The database search found 91 transcripts amongst those significantly changing in abundance were associated with interferon signalling. Transcripts were analysed using ingenuity pathway analysis to identify gene networks amongst those transcripts significantly different in EBOV compared with RESTV infection. Canonical pathway analysis was performed. Pathway alignments with a p-value of <0.05 were selected and those with high and low activation states were identified.

IL-6 Signalling was rated highly in EBOV and RESTV (z-score of 3.212 and 3.55 respectively). As was TREM1 Signalling (3.13/3.657), IL17 Signalling (2.84/3) and the acute phase response (1.859/2.694). Inflammasome pathway signalling was higher in RESTV (z-score of 2.1) than in EBOV infection (1.6) NRF2 and ERK5 signalling were both in the top10 most activated pathways for EBOV infection (1.96 and 1.8 respectively) but were not in the top10 pathways for RESTV infection. Toll-like receptor signalling was similar for both EBOV and RESTV infection (~1.8 z-score). Pathways showing the most decreased activation state (negative z-score) included Ephrin B Signalling, PPAR Signalling, Intrinsic Prothrombin Signalling and Coagulation System. Patterns of decreasing pathway activation were similar in both EBOV and RESTV infection. This analysis was repeated for transcripts showing different levels of abundance in EBOV infection compared with RESTV infection. 19 pathways were identified in this analysis. With the most significant pathways (-log10 p-value) being Interferon, Ephrin B, gamma-glutamyl, and PCP signalling.

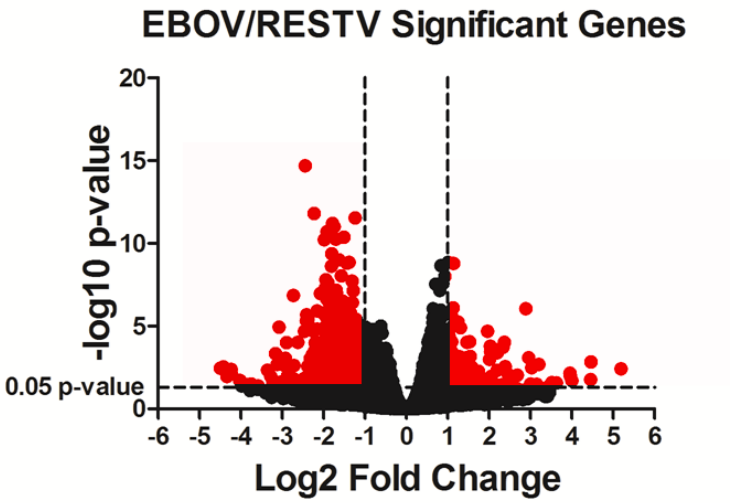


Figure 5.12: Significantly different transcripts EBOV/RESTV. This plot shows the number of transcripts with greater than +2 or -2 FC (-1/+1 log2FC) and a p-value of 0.05 (-log10 p-value 1.3). Transcripts exceeding these threshold values are coloured red, those which do not exceed these threshold values are deemed insignificant and coloured black. 484 transcripts appeared reduced in abundance compared in EBOV compared with RESTV, while 157 transcripts appear increased.

Significantly changing transcripts comparing EBOV with RESTV were also assessed for common direct interactions. Figure 5.13 shows only those transcripts with experimentally verified direct interactions. Genes with the greatest numbers of interactions in this network were HDAC3, ESR1, STAT1, NUPR1 and HSF1, all of which are known to localise to the nucleus and are involved in transcriptional regulation of genes involved in survival, activation and the immune response.

Significant gene transcripts with a fold change of >2 or less than -2 and an FDR of <0.05 when comparing EBOV with RESTV were also analysed using the STRING 10 gene networking algorithm. This analysis identified a sub-network of genes which is shown in Figure 5.14. This subnetwork was analysed for common ontology, which identified "Response to Virus" in the gene ontology database (DAVID) with a FDR of 0.000837. The transcripts encode genes which are linked by direct or indirect interactions; those with direct interactions were shown as connecting lines. These genes include OAS2, OASL, MX1, IFIT1,2 and 3, RSAD2, IFI44 and IFI44L which were all known to be regulated by interferon induction. This is further evidence for some differences in interferon induction in EBOV compared with RESTV. All of these gene transcripts appear of higher abundance in RESTV infection compared with EBOV infection of differentiated THP-1 cells.

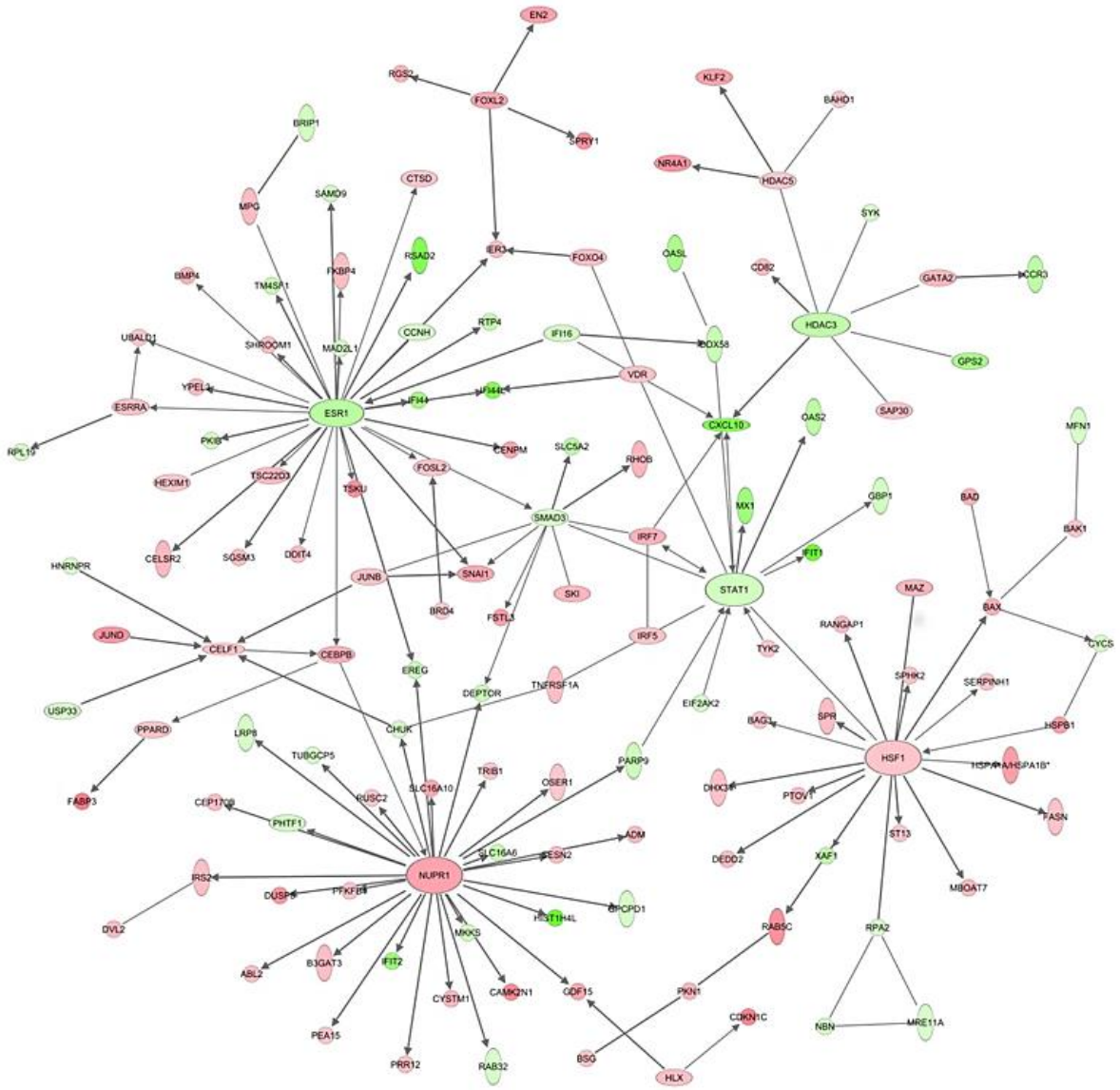


Figure 5.13: Genes with significantly changing transcript abundance were analysed using ingenuity pathway analysis (IPA) to identify pathways, regulators and experimentally verified interactions. This figure shows a network of genes with only direct interactions. An expanded figure showing all direct and indirect interactions is available in Appendix II. Changing abundance is shown by colour. Red indicated increasing transcript abundance and green indicated a decrease in transcript abundance. Only those genes identified as significant using EdgeR comparing EBOV with RESTV directly, and with an FDR of <0.05 were shown.

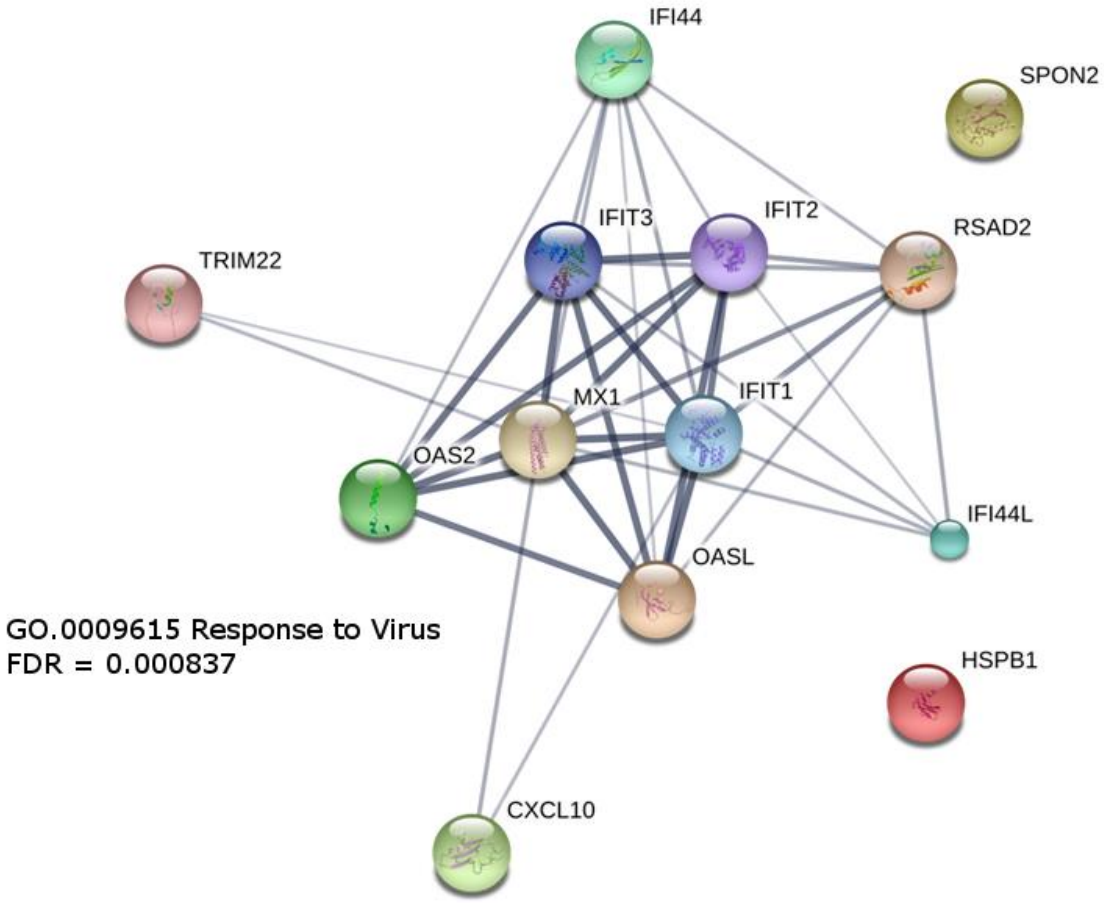


Figure 5.14: Sub-network of connected genes with significantly changing transcript abundances. Those genes with significantly increasing or decreasing abundance values comparing EBOV with RESTV were analysed in STRING 10 identifying significant over-represented ontologies. Genes with experimentally verified interactions were identified and a subnetwork identified consisting of 13 genes which aligned to the gene ontology (GO) category of "Response to Virus" with a FDR of 0.000837. Black lines indicate direct connections while grey lines indicate indirect connections.

5.3.6 Identification of differences in upstream regulation

The changes in the abundance of cellular transcripts in EBOV or RESTV infected cells compared with mock infected cells may have been caused by changes in transcription through the activation of upstream transcriptional regulators. To investigate this, IPA was used to group the changes in the abundance of transcripts and identify these regulators. To confirm the identification of upstream regulators from IPA, Site-Site Analysis was performed using the algorithm oPOSSUM v3 (219). The result of this upstream analysis is shown Figure 5.15. The analysis indicated that NFκB, IL1β, TNFα, TGFβ1 and LPS (denoting TLR4 activation) were potential major regulators of gene expression, which changed significantly in EBOV or RESTV infection compared with mock infected cells. Further evidence of increased TNF activity included the experimental FC, where in cells infected with EBOV, the TNFα transcript had a 13.28 FC in abundance compared with mock infected cells. In cells infected with RESTV, the transcript encoding TNFα increased in abundance 12.22 FC compared with mock-infected cells. TGFB1 slightly increased by 1.3 and 1.4 FC in EBOV and RESTV infection respectively; two genes downstream of TGFB1, *SQSTM1* and *MVD* were identified as having significantly changing transcript abundances. *SQSTM1* encoding Sequestosome 1 (3.15 FC in EBOV infection at 72h) involved in NFκB regulation, and *MVD*; mevulonate diphosphate decarboxylase (4.7 FC in EBOV infection at 72h), which is an important protein in cholesterol biosynthesis. IL1β was increased by 6.74 FC in EBOV infection, and 8.4 in RESTV infection. Though displaying a slightly lower z-score of 4.8 (EBOV) and 4.2 (RESTV), IL6 expression was increased by 19 FC in EBOV infection and 17 FC in RESTV infection. IL15, a potent immunomodulatory cytokine,

and HIF1A, a response gene to cell stress also were calculated to have a high z-score.

Notably genes regulated by FOS and JUN, principal components of the AP-1 pro-inflammatory complex, were over-represented in the data as well (EBOV z-score 2.2 and RESTV z-score 2.4) and FOS (EBOV z-score 2.0 and RESTV z-score 2.4). This is of particular interest as AP-1 is a potent mediator of the pro-inflammatory response, and was identified as being a significant upstream regulator of transcripts changing in EBOV compared with mock infection by Single-Site analysis using oPOSSUM (z-score: 5.392 fisher score: 11.988). The significance of AP-1 in RESTV infection appeared to be lower (z-score: 8.964 fisher score: 5.613) however the activation state appeared higher. Transcript abundance of FOS and JUN were measured, with FOS showing >2 FC transcript abundance change at all time points for both EBOV and RESTV. JUN however did not appear to change significantly throughout infection.

Overall 2669 upstream regulators were identified as being common to both EBOV and RESTV infection. 355 significant upstream regulators were unique to EBOV infection which aligned to PI3K-Akt, Ras Signalling, Steroid Biosynthesis and TNF Signalling KEGG pathways (p=<0.001). 4 of the most significant upstream regulators; those showing more or less than 2 z-score, and greater than 9 –log10 transformed p-value (IFNA2, IFNL1, IRF5, IRF7), and 20 overall in the analysis results, were involved in the interferon response. This correlates with pathway analysis results, which indicated that interferon regulation was different in EBOV infection compared with RESTV infection modelled with THP-1 cells. Single-Site analysis was performed in oPOSSUM to identify transcription factors with regulatory power common to genes that changed significantly in EBOV compared with RESTV, which identified SP1 and NFκB as having significant fisher score and z-score

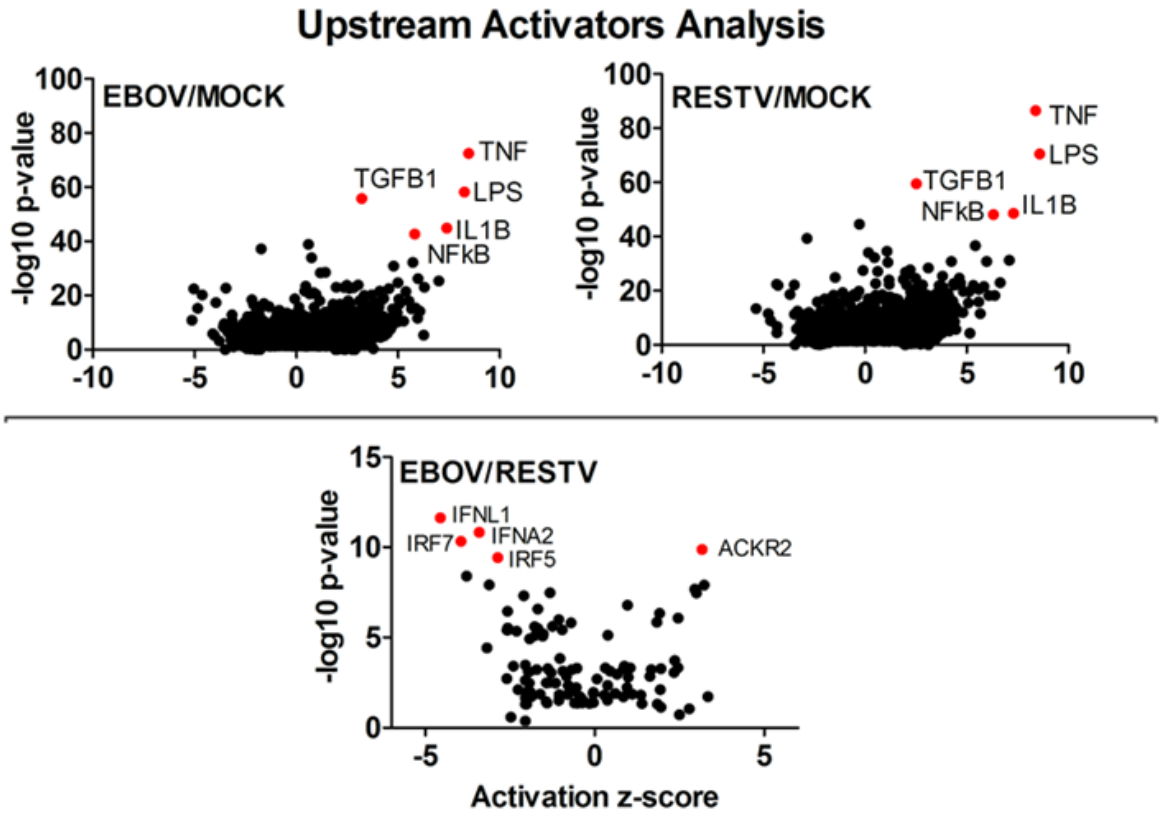


Figure 5.15: Upstream regulator analysis was performed using canonical pathway analysis data. Transcripts were aligned to canonical pathways and their cumulative increasing or decreasing abundance measurements compared with mock infection, or comparing EBOV to RESTV directly were used to calculate pathway activation score (z-score). Activation scores are ranging from decreasing activity (-10) to increasing activity (+10) based on predictive calculations. For EBOV/RESTV comparison the range is smaller (between -5 and +5) due to reduced variation in pathway activity. Only the top 5 pathways either increasing or decreasing are indicated based on both z-score and -log10 p-value. As 72h represented the most divergent time point comparing EBOV with RESTV, only data collected from this time point was used in these calculations. TNF, LPS (TLR4), TGFB1, IL1B and NFκB are shown as highly active upstream regulators in both EBOV and RESTV infection compared with mock infection. Comparing EBOV directly with RESTV shows that IFNL1, IFNA2, IRF7, IRF5 and ACKR2 are significant upstream regulators. The direct comparative analysis shows significant though reduced -log10 p-value due to reduced numbers of transcripts.

5.4 Validation of RNASeq data.

Classical methods of measuring transcript abundance are still critical in high throughput transcriptomic studies using microarrays or RNASeq. qRT-PCR has been used for many years to profile the relative abundance of transcripts. Comparing the patterns of relative abundance measured by qRT-PCR with RNASeq transcript counts allows validation of the calculated fold change in selected transcripts. This was particularly important due to the lengthy and complex bioinformatics analysis required to reliably calculate transcript abundance using RNASeq data, which may introduce unexpected errors or bias. Testing all transcripts in an RNASeq database was impractical; which may include all transcripts aligning to all regions of the genome. Instead, a qRT-PCR targeting antiviral genes was used to further explore the effects of EBOV and RESTV on the antiviral response in differentiated THP-1 cells.

5.4.1 Infection of A549, THP-1 and HepG2 cells

The possibility remained that patterns of transcript abundance measured by RNASeq were unique to THP-1 cells. To investigate whether these patterns of gene induction were similar in other cell types permissive for *ebolavirus* infection, A549 and HEPG2 cells were infected in parallel with differentiated THP-1 cells.

5.4.2 Replication kinetics of EBOV and RESTV in three human cell lines

RNASeq and quantitative proteomics was used to analyse the host response to EBOV and RESTV infection in THP-1 cells. However, first, to test whether the growth of the two viruses was potentially equivalent (and allowing a direct comparison of the host response to be made), the replication of the viruses in cell culture and the amount of virus released was assayed by measuring the abundance of the viral genome as a proxy for this. Duplicate infections were performed at an MOI of 5 for the two different viruses in the three different cell lines used in this study (THP1, A549 and HEPG2).

qRT-PCR was used to measure viral genome abundance at 0h 6h, 24h, 48h and 72h post-infection. Equivalent viral genome copies were calculated by comparing the abundance of the viral genome with a synthetically generated control RNA equivalent to the GP gene of EBOV (Makona variant) or RESTV. The data was generated as genome equivalents and presented as plots (*Figure 5.16*). Statistical analysis by two-way ANOVA showed that EBOV and RESTV viral genome abundance was equivalent in total RNA purified from cells and in RNA purified from the supernatants for the THP-1 and HEPG2 cell lines at 72h post-infection.

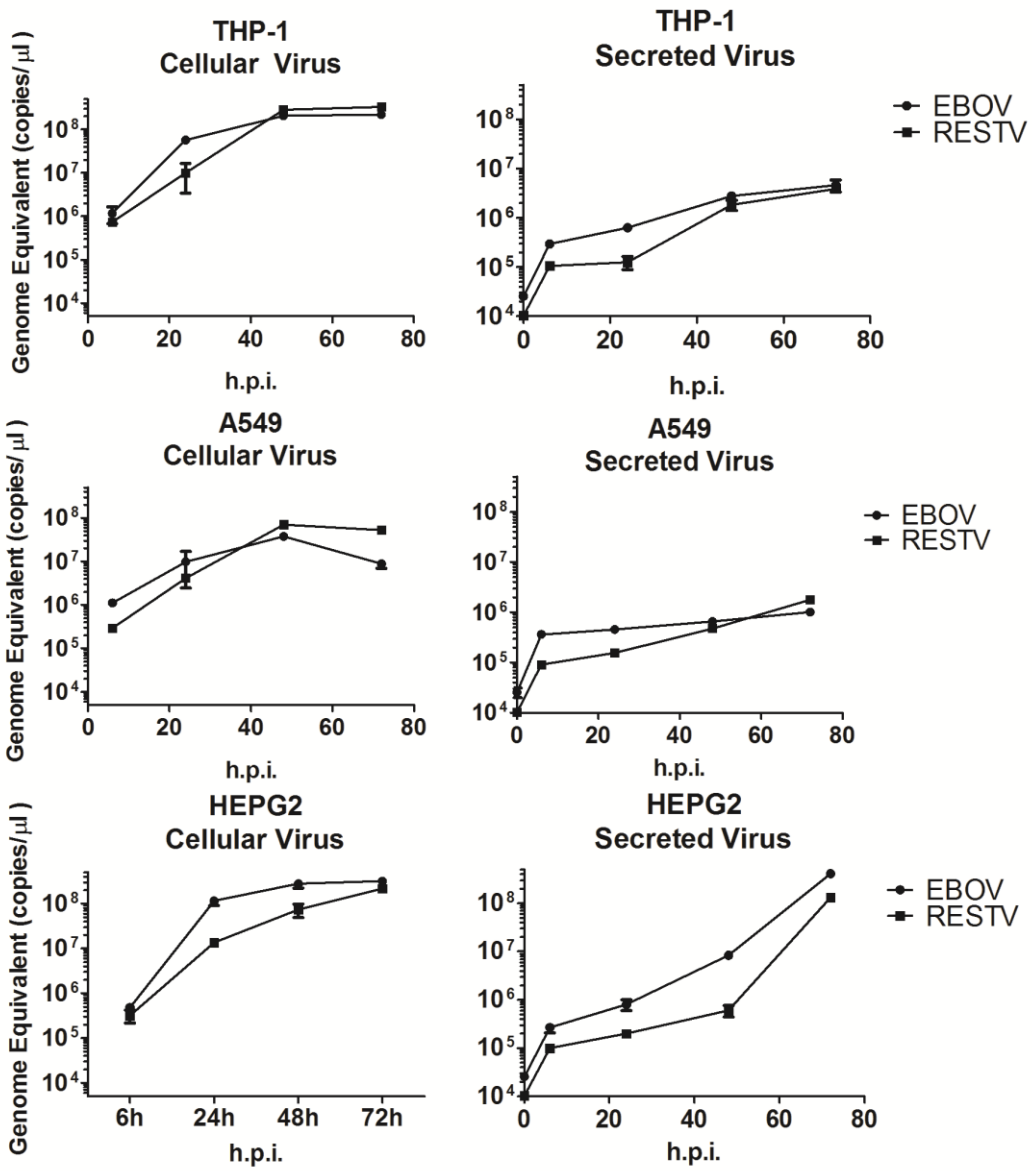


Figure 5.16: Genome abundance of EBOV and RESTV from THP-1, A549 and HEPG12 cells from 0h.p.i. to 72h.p.i measured in supernatant or from 6h.p.i to 72h.p.i measured in cells using qRT-PCR. The assay was directed to detect mRNA corresponding to the GP gene of either EBOV or RESTV and quantified using the standard curve method with an artificially synthesized RNA transcript control. Data shown is in the log10 scale to aid visualization. Data based on 2 biological replicates for each condition and PCR performed with 3 technical replicates per condition for accurate measurement of viral RNA.

5.4.3 qRT-PCR analysis validates RNASeq abundance measurements

Before an in depth analysis of the RNASeq dataset was performed the expression profile of selected transcripts was validated by qRT-PCR. Here, the transcript abundances of 84 antiviral genes were profiled by qRT-PCR and compared with calculated gene transcript counts acquired by RNASeq. A full list of antiviral genes measured in this assay is available in subsection 2.3.4. Further to confirming the analysis in THP-1 cells, the transcription profile was also investigated in A549 and HEPG2. Total RNA was collected at 72h and the abundance of transcripts from selected genes measured by qRT-PCR (Figure 5.17). Similar to the RNASeq data, principle component analysis suggested that the abundance of transcripts in EBOV and RESTV infected cells were generally similar compared with mock-infected cells (Figure 5.17A). The abundance of transcripts encoding antiviral proteins in THP-1 and HEPG2 cell lines closely correlated closely, whereas the profile in A549 infections was different. Figure 5.17B shows a comparison of relative expression data collected for each cell line. This analysis shows a close correlation in antiviral gene transcription between EBOV and RESTV infected THP-1, A549 and HEPG2 cells, indicating that the patterns of transcription of antiviral response genes are similar across all human cell lines tested. Comparison of fold change in abundance of transcripts measured by RNASeq and qRT-PCR was performed using a Spearman correlation analysis. This analysis demonstrated that RNASeq and qRT-PCR measured transcript abundance correlated closely, helping to validate the RNASeq dataset Figure 5.17C and D.

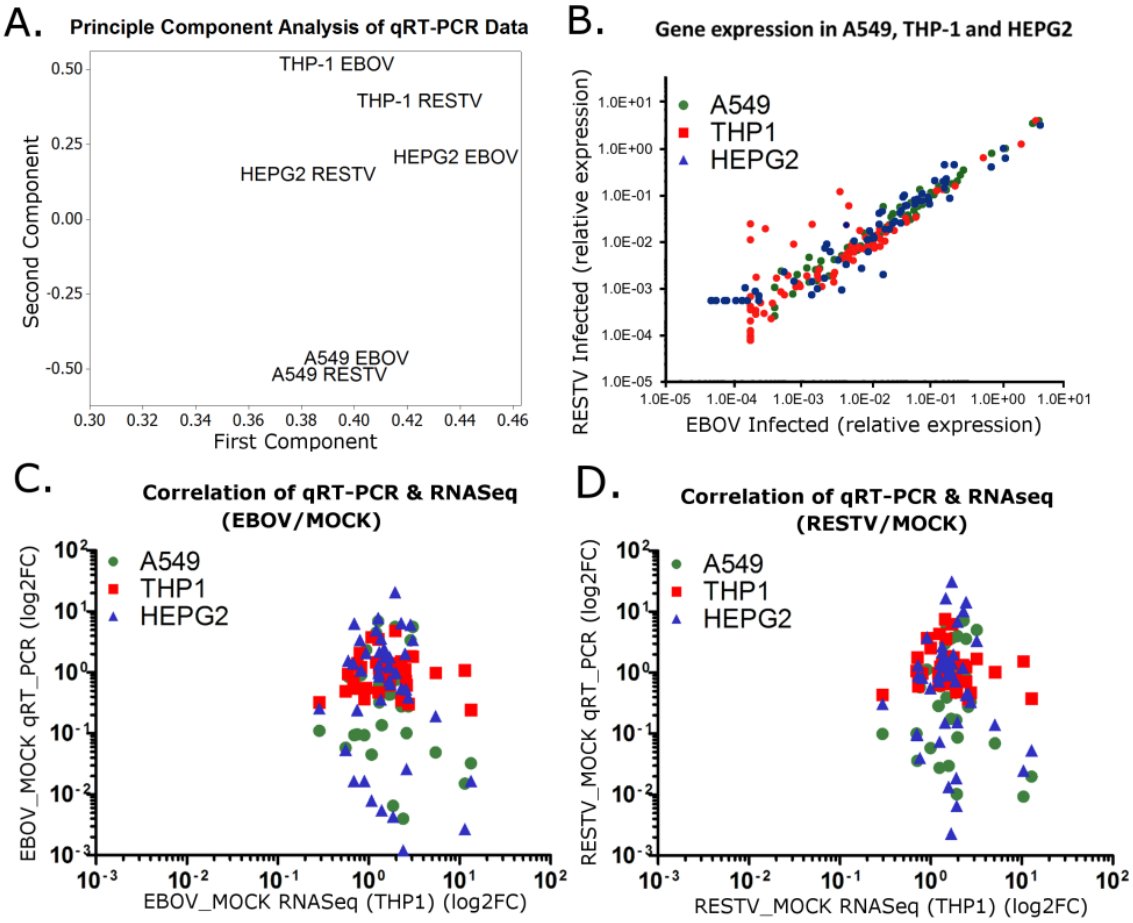


Figure 5.17: qRT-PCR analysis of 84 antiviral transcripts in THP-1, A549 and HEPG2 cells infected with either EBOV or RESTV. (A) A principle component analysis plot showing the correlation and relative distance contrasting each condition, EBOV and RESTV conditions cluster by cell type, with the most divergent cell type appearing to be A549 cells. THP-1 and HEPG2 both show more diversity comparing EBOV with RESTV than in A549 cells. (B) Correlation of transcript abundance in A549, HEPG2 and THP-1 cell lines between EBOV and RESTV induced gene expression. This analysis shows that most genes correlate well, with more diversity seen in THP-1 cell lines. (C and D) This analysis compares transcript abundance measured by qRT-PCR with measurements made with RNASeq. This evaluation of RNASeq data validity using a more routine method in several different cell lines demonstrates that RNASeq measurements are comparable with qRT-PCR values in all three cell lines for this panel of 84 antiviral genes tested.

5.4.4 Interferon stimulated transcript abundance were decreased early in infection while inflammation associated transcript abundance were increased.

Expectation was that EBOV and RESTV infection of THP-1 derived macrophages would result in a measurable increase in transcript encoding interferon response genes. Transcriptomic analysis revealed that the interferon response in THP-1 cells infected with EBOV and RESTV did not increase as expected, and some transcripts decreased early in infection. Cellular RNA samples collected at 6h, 24h, 48h and 72h post infection were analysed by qRT-PCR to measure the abundance of transcripts encoding antiviral genes. An analysis of the correlation in transcript abundance measured in EBOV and RESTV infected THP-1 cells showed that data collected at 6h and 24h correlated better (R^2 >0.5) than at 48h and 72h. This analysis (

Figure 5.18) revealed the same patterns of correlation as had been acquired from the first infection experiment analysed by RNASeq, and indicated that transcript abundance at later time points were more divergent. A heatmap of the antiviral genes was generated to visualise the fold change in transcript abundance across all time points. Interferon stimulated transcripts showed the greatest abundance in RESTV infected cells at 72h, though FC was only slightly above threshold (>2). The attenuated interferon stimulation was observed throughout the time course including at 6h post infection, persisting at all other time points. Conversely, transcripts involved in the inflammatory response (CCL5, CXCL8, RELA, TNF, NFKB1) show greater abundance at 6h post infection in both EBOV and RESTV, compared with the other time points (Figure 5.19)

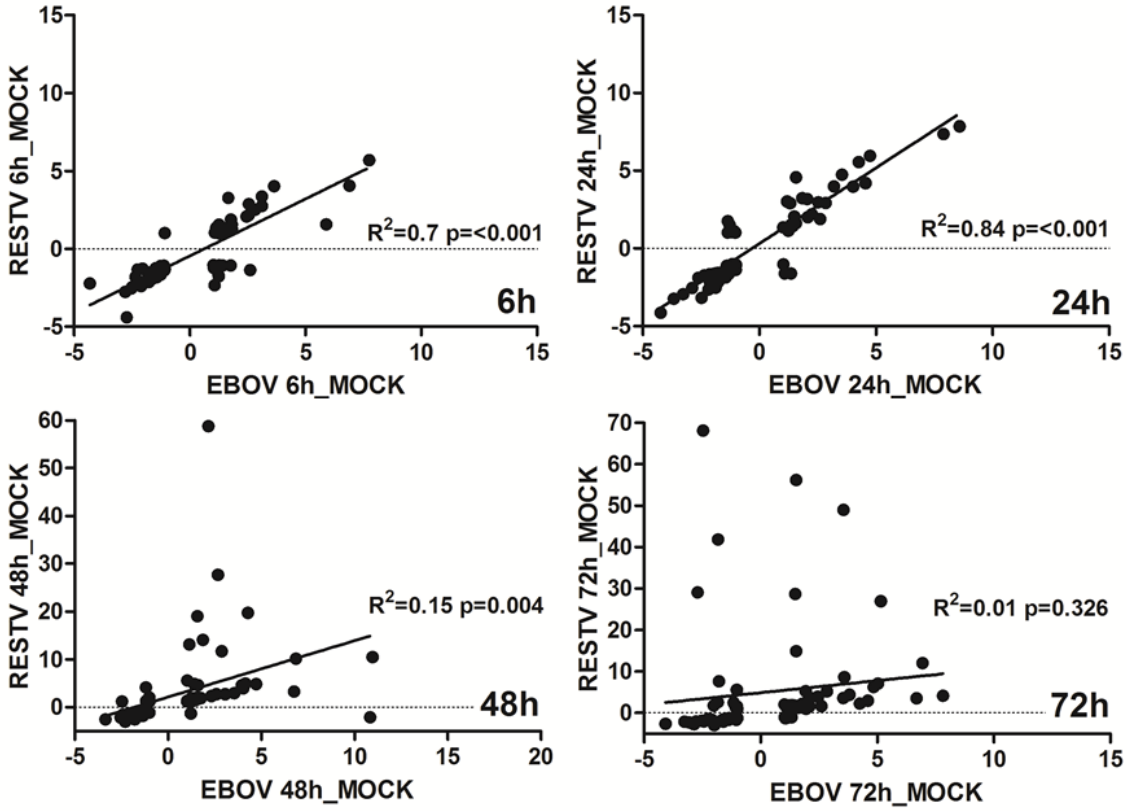


Figure 5.18: Spearman analysis of the correlation between qRT-PCR data obtained from THP-1 cells infected with RESTV or EBOV at 6h, 24h, 48h or 72h. R^2 values were given to indicate correlation. An R^2 value of 1 indicated a perfect correlation, and 0 indicated no correlation. Significance testing was performed by Spearman's correlation without correction, where a p-value of <0.05 was likely to be a significant test of correlation. All transcript targets in the qRT-PCR arrays were utilised in this analysis. X axes show fold change in qRT-PCR delta-cT values for EBOV infected THP-1 cells. Y axes show fold change in qRT-PCR delta-cT values for RESTV infected THP-1 cells.

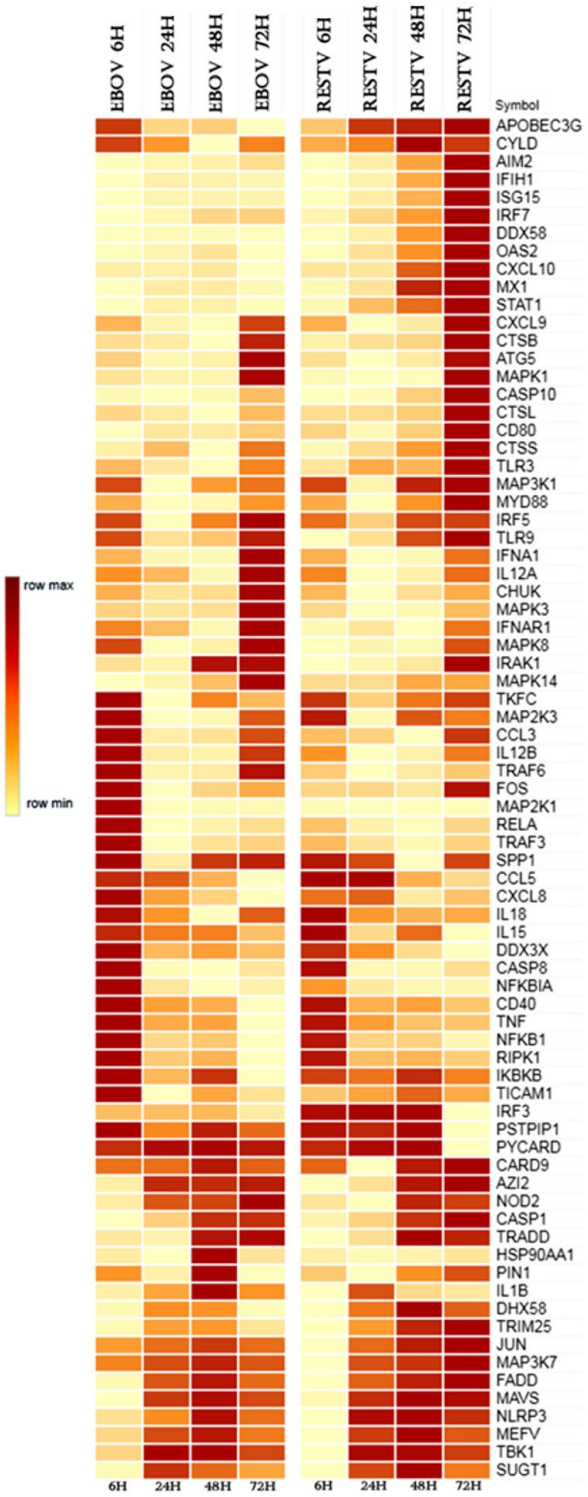


Figure 5.19: Heatmap of fold change in qRT-PCR delta-cT values produced using MORPHEUS (Broad Institute). Heatmaps are divided into EBOV (left) and RESTV (right) datasets. Data was arranged sequentially as a time course from 6h, 24h, 48h and 72h post infection. Each row is coloured on a scale of the maximum value in a row (Dark Red) to the minimum value in the row (Pale Yellow). Pearson hierarchal clustering was performed.

5.5 Conclusions

This chapter describes work done to identify elements of the host response unique to the Makona variant of EBOV in a macrophage cell model. Transcriptomic analysis has revealed a network of genes where transcript abundance appeared different in EBOV infection compared with RESTV infection. The function of these genes was mostly related to antiviral response and interferon signalling. Common traits appeared to be a potent activation of macrophage-associated pro-inflammatory cytokines, many of which have been implicated in EVD. These potent reactions to viral infection appeared shared by both EBOV and RESTV. Several key upstream regulators were identified, including NFκB, TNF and TLR4 and AP1. The transcription factors NFκB and SP1 were identified as having significant roles in the regulation of transcripts which appeared different in EBOV compared with RESTV.

Differences and similarities in the pro-inflammatory response induced by Makona and RESTV have been further explored in Chapter 6. Proteomic analysis was performed using SILAC labelled differentiated THP-1 cells, to ascertain whether the host response described with transcriptomics was reflected at the protein level, and to investigate the role of key host regulators in the life cycle of *ebolaviruses*.

6 Proteomic analysis of EBOV and RESTV infected THP-1 derived macrophages

6.1 Introduction

Transcriptomic analysis revealed that most discernible differences in the host response to EBOV when compared with RESTV occurred in the antiviral immune response. Of particular interest was the increased activity measured in response genes controlled by the pro-inflammatory regulators NFκB, TNFα and IL1β. To produce further evidence that EBOV and RESTV infections result in the activation of these pro-inflammatory regulators, mass spectrometry was used to probe the proteomic changes taking place during infection. Proteomic data was also compared with transcriptomic analyses to identify discrepancies in the abundance of gene products at the transcript and protein level and to ensure transcript abundance data was otherwise reflected at the protein level. Models of EBOV and RESTV lifecycle were created and used to assess whether inhibiting key host regulators may affect virus lifecycle.

6.6.1 Chapter aims

In Chapter 5 evidence was presented that both RESTV and EBOV infection of THP-1 macrophages elicited high levels of pro-inflammatory gene transcription and with Makona infection producing measurably greater transcriptional induction of pro-inflammatory regulators than RESTV. The work also illustrated that some differences may exist in the antiviral response, in contrast to that observed in A549 cells, which required further study. In this chapter transcriptomic differences were validated using proteomics, confirming the increased activity of pro-inflammatory regulators and suggesting a key role for these regulators in the lifecycle of the EBOV, Makona variant.

6.2 Chapter Methods

6.2.1 Processing of samples for Proteomic analysis

A full report on the methodology utilised in the SILAC proteomics experiment discussed in this chapter is available in Chapter 2. Several processing methods were tested. The final method used SILAC labelling, extraction of protein with the Qiagen AllPrep kit coupled to In-Gel Digestion which displayed interesting results,

6.2.2 Optimisation of protein extraction methods

In the final testing of the FASP-IGD method a sample containing unlabelled A549 cells infected with EBOV for 24h was divided into 5 aliquots. Each aliquot was then spiked with a commercially available control protein, and was processed using the AllPrep-IGD method. Peptides originating from the spiked protein control were identified in all samples and were of similar. Comparison of all peptides identified in all 5 aliquots showed low variation. A PCA was performed to evaluate the correlation, indicating low data divergence (Figure 6.1). Multivariate GLM analysis of correlation was performed. Showing that all samples were closely correlating (>0.7 in all cases ; Table 6.1).

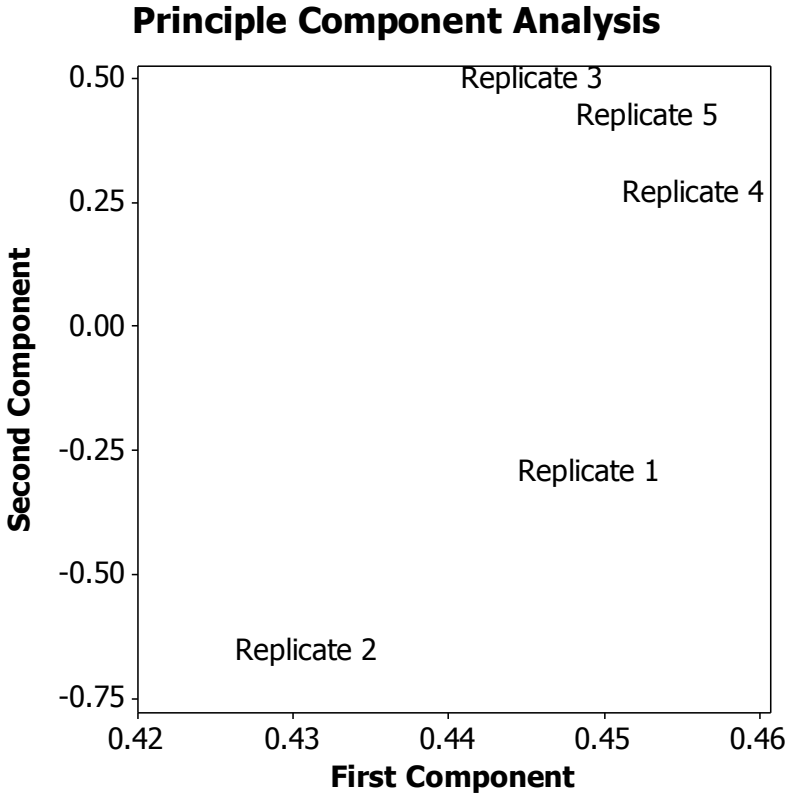


Figure 6.1: Principal Component Analysis (PCA) of peptide quantitation recovered from 5 separate identical samples of the same sample original specimen. PCA analysis shows slight clustering of 3 replicates, while 2 replicates cluster less well. The scale of this difference is small. Shown are replicate data modelling by the first and second components, indicating a minor separation by the second component. The scales shown indicate that these differences are incredibly small.

	REPLICATE 1	REPLICATE 2	REPLICATE 3	REPLICATE 4	REPLICATE 5
REPLICATE 1	1	0.833	0.85	0.833	0.804
REPLICATE 2	0.833	1	0.795	0.844	0.814
REPLICATE 3	0.85	0.795	1	0.912	0.922
REPLICATE 4	0.833	0.844	0.912	1	0.943
REPLICATE 5	0.804	0.814	0.922	0.943	1

Table 6.1: Datasets were analysed by Pearson correlation analysis to calculate the similarities between replicates and confirm principal component analysis. All replicate showed a strong correlation (>0.75). Despite this, it had been hoped that due to the small numbers of replicated available for experimentation and due to funding constraints, even a correlation of <0.8 was seen to be unsuitable as it would increase the number of label free replicates needed for analysis. Therefore unsuitable amounts of variation were introduced into the analysis in label free proteomics and required compensation when designing experiments. Thus SILAC metabolic labelling was selected as method of choice to circumvent this limitation.

6.2.3 SILAC Labelling and validation

Metabolic labelling of THP-1 and A549 cells was performed as described in chapter 2. Cells metabolically labelled with an isotope of L-lysine containing an additional carbon atom were designated “heavy” while those metabolically labelled with a native form of L-lysine were labelled as “light”. Stable native arginine was also added to media and was not used in analytical discrimination between these samples after combination. Samples were combined at a ratio of heavy: light 1:1 and processed as a single sample. SILAC was analysed for validation that the labelling had been incorporated correctly. Example peptide signatures of peptides from metabolically heavy and light cells were shown in figure 6.2

6.2.4 In-Gel digestion of peptides for ESI-MS analysis

Combined cell lysates in LDS were loaded onto a 4-12% bis-acrylamide gel and separated into discrete bands as described in chapter 2 The result of this separation is shown in Figure 6.3. Bands were excised and treated with a gradient of acetonitrile to dehydrate gel fragments. A buffer of ammonium bicarbonate containing sequencing grade Trypsin was then added and the fragments incubated for 16h at 37 degrees with rotation. Resulting peptides were eluted from fragments and treated as described in chapter 2 prior to analysis using an Orbitrap Velos mass spectrometer by MS/MS.

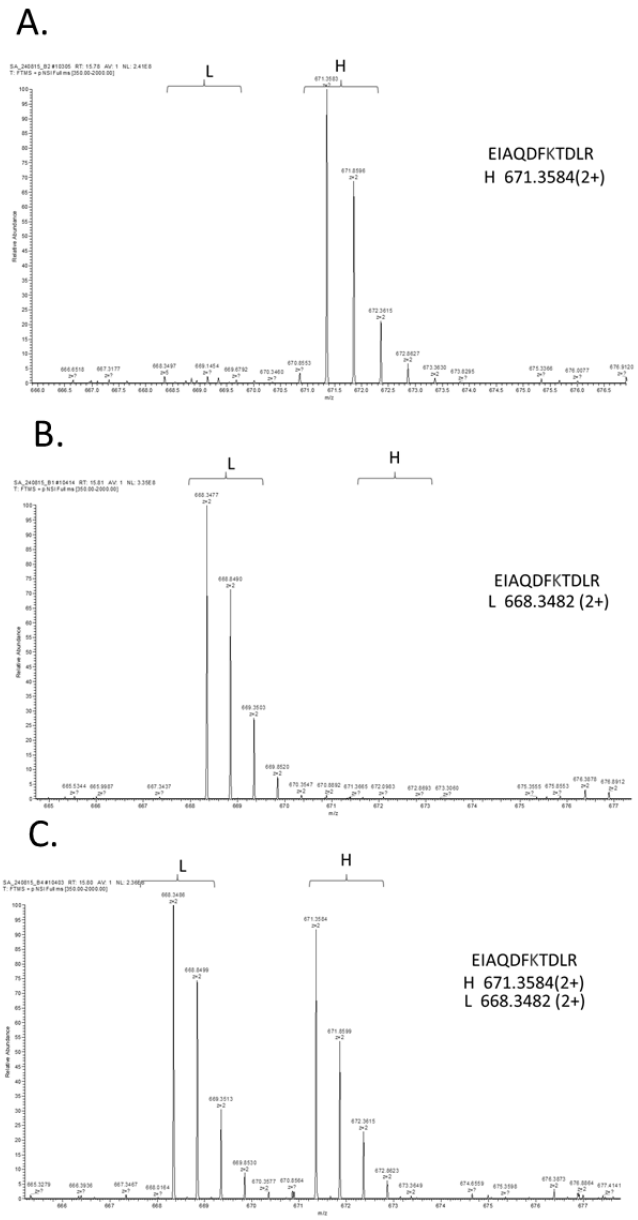


figure 6.2: Peptide signatures identified in three separate metabolically labelled samples derived from the SILAC labelling of THP-1 cells to illustrate the effectiveness of labelling. (A) peptide signature identified in a sample derived from cells which have been cultured in the presence of "heavy" lysine. Observable is indicated are the peaks of predicted mass/charge ratio for a heavy labelled peptide. The predicted charge ratio position of a comparable "light" unlabelled peak is also shown, indicating that in "Heavy" THP-1 macrophages no "light" unlabelled versions of this peptide were identified. (B) the "light" comparator for sample A, showing a peak for a peptide corresponding only to the wildtype m/c ratio. (C) Samples A and B were combined and analysed simultaneously, as a consequence both "heavy" and "light" peaks are shown.

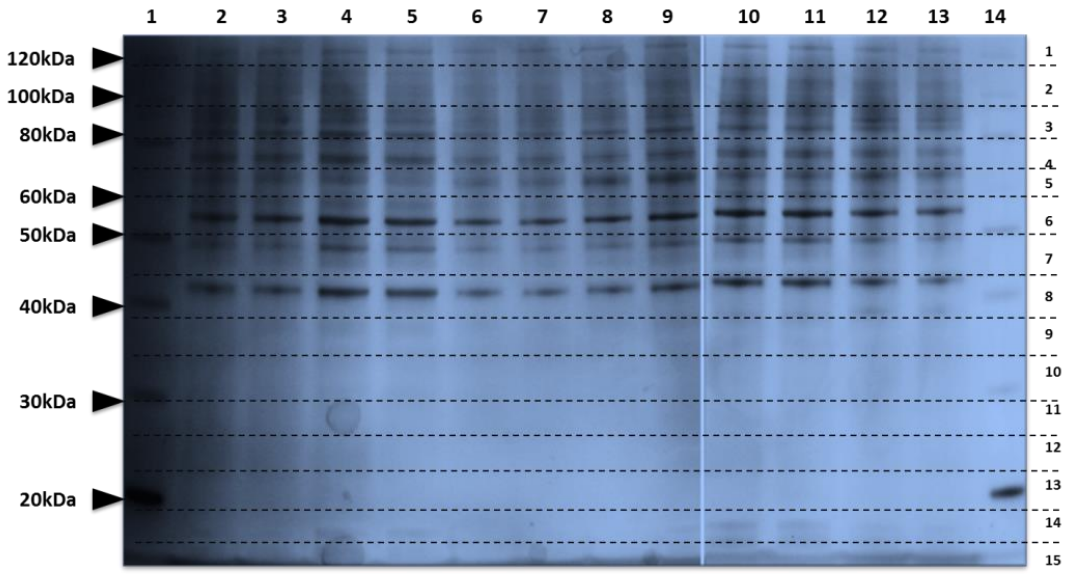


Figure 6.3: Coomassie stained 4-12% bis-acrylamide gel loaded with whole cell lysates cleaned up by FASP and suspended in denaturing LDS buffer. Shown is a compensate image of two separate gels separated by lanes 9 and 10. Lanes 1 and 14 contain pre-stained protein standards ranging from 20kDa to 120kDa. The remaining lanes were arranged as paired samples of EBOV infected THP-1 cells at 24h, 48h and 72h (lanes 2-7) followed by paired samples of RESTV infected THP-1 cells at 14h, 48h and 72h. (lanes 8-13) The gel was run for 50 minutes on a vertical gel tank at 200V and imaged using a GeneSYS gel docking system.

6.2.5 Designing the mini-genome system

Full details of the trVLP (transcription-replicon competent Virus Like Particle) and mini-replicon assay methodology is available in chapter 2. The VLP component of the assay was developed *de novo* using sequence data acquired and described in chapter 3, pertaining to the Makona variant of EBOV. Sequencing data acquired from the analysis of RESTV infection in A549 cells was used to generate sequences encoding the VP40, GP and VP24 in the pUC57_A388 plasmid vector system. These were added to existing constructs encoding the L, VP30, VP35, NP and Mini-Replicon. The vector and inserts encoding all proteins in the assay were synthesised de novo by GeneART, part of Thermo Scientific; as described chapter 2. All protein coding sequences were codon optimised for production in human cell lines, as it was advised that this may improve viral protein yield from transfections.

The mini-replicon for the model system of the Makona variant of EBOV was originally designed by Dr. Weining Wu of University of Liverpool. The mini-replicon comprises a renilla luciferase coding sequence, flanked by the ultimate 5' and 3' terminal regions of the EBOV genome. This design was also used to construct a mini-replicon for the RESTV model system.

Constructs were sequenced using illumina sequencing to verify sequence integrity prior to plasmid amplification and use in the HEK-293T cell line. The trVLP and mini-genome systems are useful for the modelling of virus replication, cell entry and cell egress, without the need to use live virus, which presents a significant biosafety hazard and is technically challenging to assay. The system utilised in this study was based on the original designs created by Dr. Thomas Hoenen of the Centre of Disease Control (72).

The plasmid components were transfected into HEK-293T cells or other appropriate cell lines, alongside a T7 polymerase encoded on an additional plasmid. The plasmids produce the viral proteins necessary for replication and encapsidation. Mini-replicon derived mRNA is bound by the viral transcriptional machinery and replicated, mimicking the processes observed during replication by true EBOV and RESTV. The proportion of translated luciferase from mini-replicon mRNA is proportional to the activity of the viral replication machinery, and is assayed by measuring luminescence of transfected cells. To monitor the efficiency of viral entry or egress the VLP component may be added to the mini-replicon assay, which allows transition of mini-replicon and viral ribonucleoprotein complexes from transfected cells to naïve cells, modelling the infection processes. A detailed schematic diagram of the mini-genome system is available in Figure 6.4.

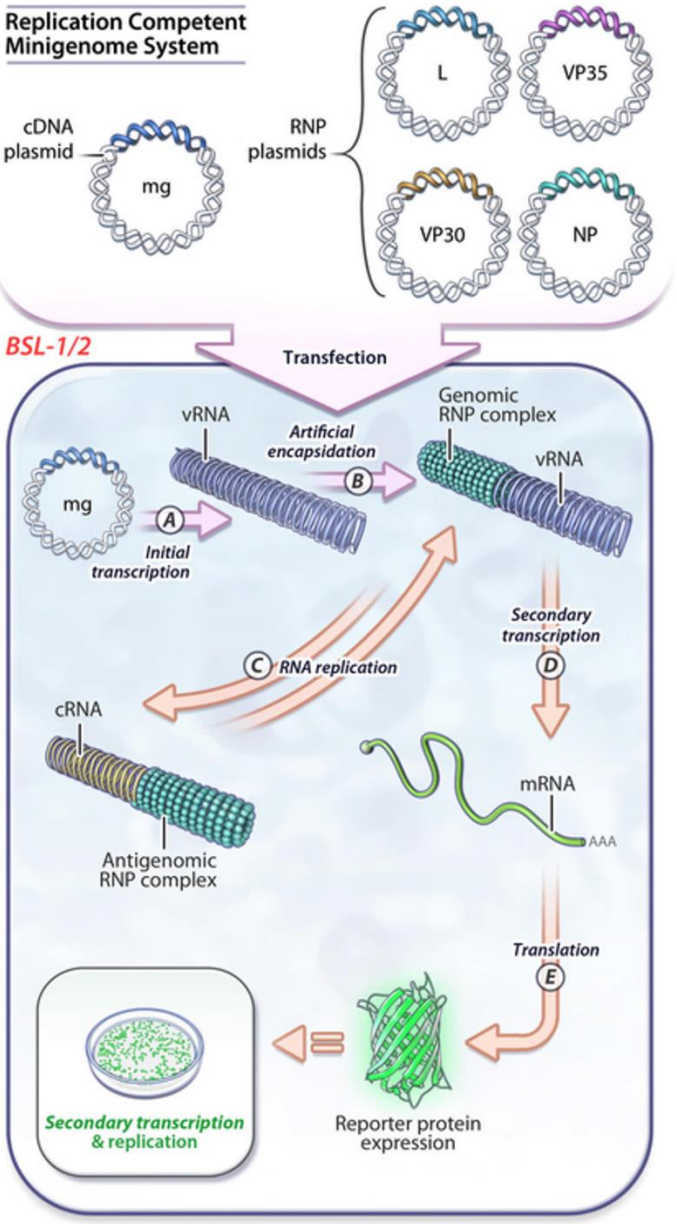


Figure 6.4: Schematic representation of the processes involved in the generation of an ebolavirus mini-genome system. The model system reflects the processes which take place during a replication cycle of EBOV or RESTV. The plasmids encoding the mini-genome were shown (VP30, VP35, NP and L) alongside the minigenome. The containment requirements were also highlighted; the system permits the study of ebolavirus lifecycle under BSL2/CL2 conditions. This diagram was adapted from a review by Hoenen et al. (72)

6.2.6 Plasmid sequence verification by next-generation sequencing analysis.

All plasmid constructs were sequenced using a MISeq prior to use in the mini-replicon or trVLP assays. As the insertion site sequences were determined using Sanger sequencing by GeneART it was deemed unnecessary to sequence the plasmids individually. Instead all plasmids were pooled into a single sample and sequenced with a single indexing barcode. The sequences for each plasmid encoded viral protein were then de-convoluted bioinformatically by mapping sample sequences directly to EBOV or RESTV codon optimised sequences from the plasmids original construction. This analysis showed that no point mutations were detected and full sequences for all proteins were identified.

6.2.7 Testing with the trVLP in BSRT7 cells and HEK-293T

Initial testing of the trVLP in BSRT7 cells (constitutively expressing T7 polymerase) and HEK-293T cells showed high luciferase activity, however results of trVLP transmission assays were highly variable and were not used in determining the significance of results in this study. Additionally storage of trVLPs at any temperature for longer than 24 hours resulted in loss of function, and this was deemed unsuitable to allow consistent results for large scale screening to be performed using this method. The mini-genome assay alone without transmissible components was trialled in BSRT7 cells, displaying high levels of luminescence, which was also later found in HEK-293T cells. The differences from control (omitting the L polymerase and thus expected to be negative for luminescence) were great enough to allow a suitable dynamic range for testing. Thus the mini-replicon was used in later experiments to model the effects of inhibitors on EBOV and RESTV lifecycle.

6.2.8 Cytotoxicity testing of drugs and mini-replicon in cell culture

To evaluate the negative cell effects of transfecting multiple plasmids encoding the EBOV and RESTV viral proteins cytotoxicity assays were performed. Additionally to identify the optimum concentrations of drugs for inhibition of host factors without detrimental effects on cell viability, an MTT assay was performed. A full description of the cytotoxicity assays used in this project is available in chapter 2.

In order to evaluate the cytotoxicity of drugs and mini-replicon a non-radiological fluorescence assay was used, known as the CellTitre assay (Promega). Production of fluorescent dye correlates with the level of cell viability detected. To confirm the results of this assay an MTT assay was used. Cell viability was determined by measuring detectable changes in colourimetric changes through the metabolism of MTT into formazan. This colourimetric change correlates with the level of cell viability.

6.3 Proteomic analysis of EBOV and RESTV infection of THP-1 cells

6.3.1 Peptide identification and differential analysis

Proteomics analysis identified between ~1100 unique proteins in samples from EBOV and RESTV. Quantitation was performed and fold change calculated comparing samples of infected cells with mock infected to identify significant changes in protein abundance. Proteomics data from EBOV infected cells was compared with data acquired from RESTV infected cells. <10% of the identified protein signatures showed significant differences between RESTV and EBOV at 24h, 48h and 72h post infection. Pearson correlation analysis was performed illustrating a high degree of correlation between proteomic data acquired from RESTV and EBOV infected cells (>0.8 p=<0.001). Proteomic data from EBOV and RESTV infected cells analysed at 24h, 48h and 72h were compared with corresponding transcript abundance measurements acquired through NGS. Pearson's correlation analysis revealed a reasonable correlation of proteomic and transcriptomic abundance measurements in RESTV infected cells (0.65 p=<0.001), while the correlation of data from EBOV infected cells was revealed to be less strong (0.54 p=<0.001). The results of this analysis were shown in Figure 6.5. Together this analysis reveals that most proteins identified in proteomics data did not change significantly in EBOV or RESTV infection which correlates with transcript abundance data in around half of cases.

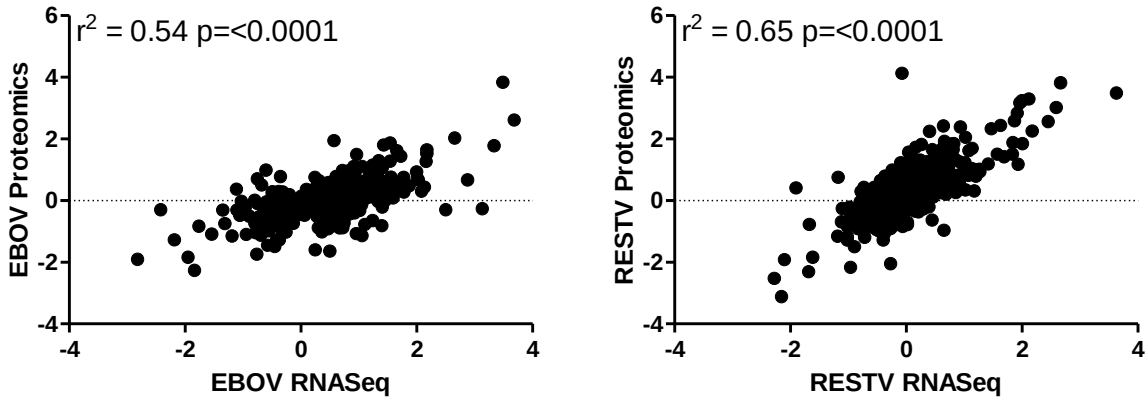


Figure 6.5: Correlation between transcript abundance data and protein abundance data in EBOV and RESTV infected cells. Shown is data amalgamated from all three time points assayed (24h, 48h and 72h post infection). The correlation was interrogated by Pearson correlation, indicating an r squared value; where this value is greater than 0.5 it indicated a positive correlation between these sample datasets. The confidence in this analysis is indicated by p-value. Each data point shown illustrates a particular transcript and corresponding protein abundance measurements at either 24h, 48h or 72h. Separate graphs were shown for data acquired from EBOV or RESTV infected THP-1 cells.

6.3.2 EBOV and RESTV do not inhibit global protein production in infected cells

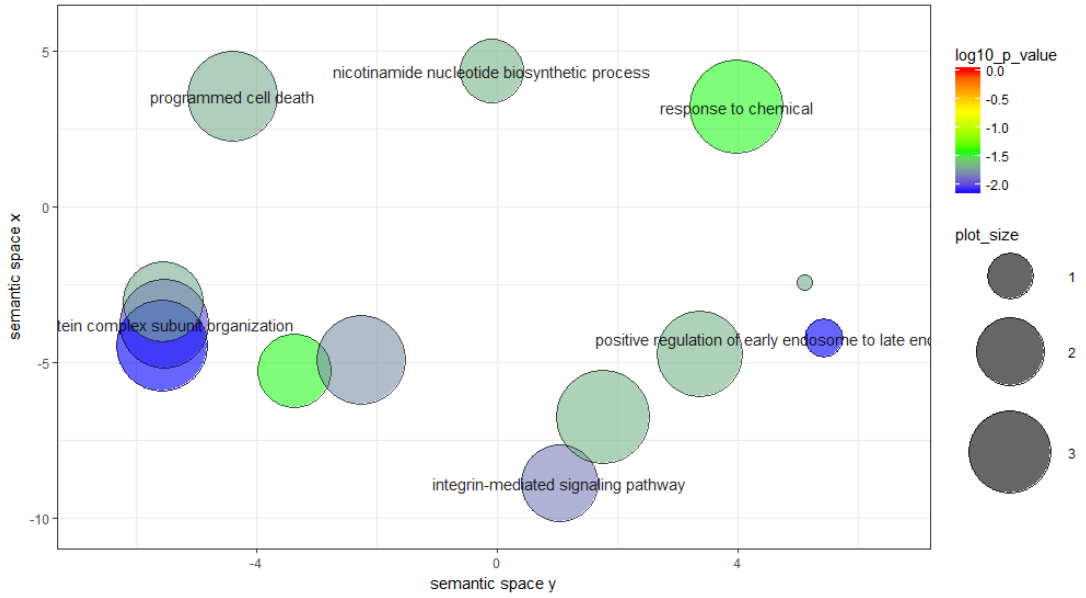
Based on results of proteomic analysis few proteins observed change significantly during EBOV or RESTV infection at 24h, 48h or 72h. This combined with initial preliminary protein quantitation and gel staining of whole cell lysate does not indicate a global decrease in protein production as is observed with other viruses.

4.5.7 Identification of significant protein functional categories.

Investigating whether the proteins in this analysis belonged to an over-represented group was important to ascertain whether this group was relevant to the biology being tested, or whether it may indicate a technical bias in protein acquisition.

Proteins which showed significant changes in abundance were annotated using GProX and functional categories trimmed using Revigo. A bubble plot showing the significance of the identified functional categories was created in the R environment. The results of this analysis were shown in Figure 6.6. The most significant functional category identified in proteomic analysis of EBOV and RESTV infected THP-1 cells was response to endogenous and exogenous stimuli. Additionally, those functional categories highlighted in analysis of EBOV infected cells were positive regulation of endosomes, integrin activity and protein biosynthetic pathways. Functional categories highlighted in proteomic analysis of RESTV infected cells included sertoli cell barrier activation, negative regulation of apoptotic pathways, control of cellular compartment arrangement, T-cell antigen processing and nicotinamide nucleotide biosynthesis. Overall the analysis indicated that a wide range of proteins were identifiable with relationships to varied parts of the cell and sub-cellular compartments, no intrinsic bias was observed.

A.



B.

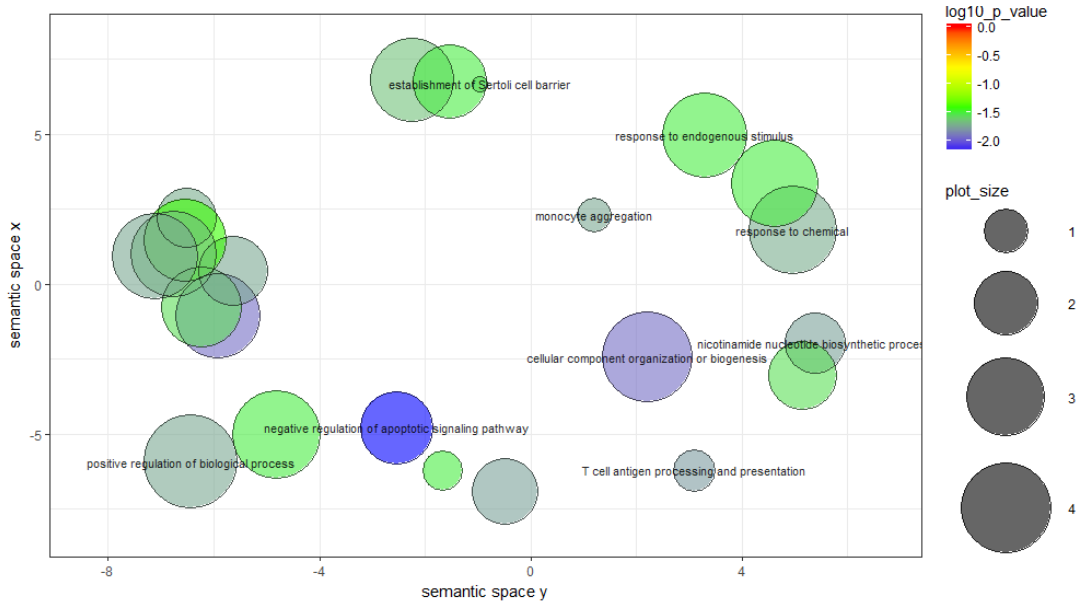


Figure 6.6: Ontological analysis of proteomic data from (A) EBOV or (B) RESTV infected THP-1 derived macrophages. Data shown pertains only to those proteins demonstrating a significant increase or decrease at one or more time points assayed (24h, 48h or 72h). Categories were merged in REVIGO. The bubbleplot was generated using an R-script generated in REVIGO and processed in the R environment. The size of each bubble represents the relative number of proteins which have been assigned to that category. A scale bar showing Red (0.0) to Blue (2.0) is provided as an indicator of significance of each categorisation.

6.3.3 KEGG Pathway analysis of combined proteomic and transcriptomic analysis

The correlation of proteomic and transcriptomic data made it likely that those proteins with associated changes at the transcript abundance level may represent a significant subset for further analysis. Additionally it was important to ascertain whether there was a biased over-representation of transcripts associated with a particular functional category where both transcriptional and proteomic data was available.

Protein abundance data was filtered to only include data for proteins for which there was corresponding transcript abundance data. Those proteins which showed a significant increase or decrease in abundance were analysed using the KEGG ontology database. This analysis identified 23 significantly represented functional categories. The relative significance of this alignment was compared between EBOV and RESTV to indicate the level of confidence associated with the number of proteins fitting into a particular category. 12 of these KEGG pathways were associated with inflammation or infection. Of particular note was the identification of TNF signalling pathway, PI3K-Akt signalling and NFκB signalling in this analysis The associated -log10 p-value for these relevant pathways means they were represented with greater significance in RESTV compared with EBOV infection associated data. A graph summarising this analysis is shown in Figure 6.7.

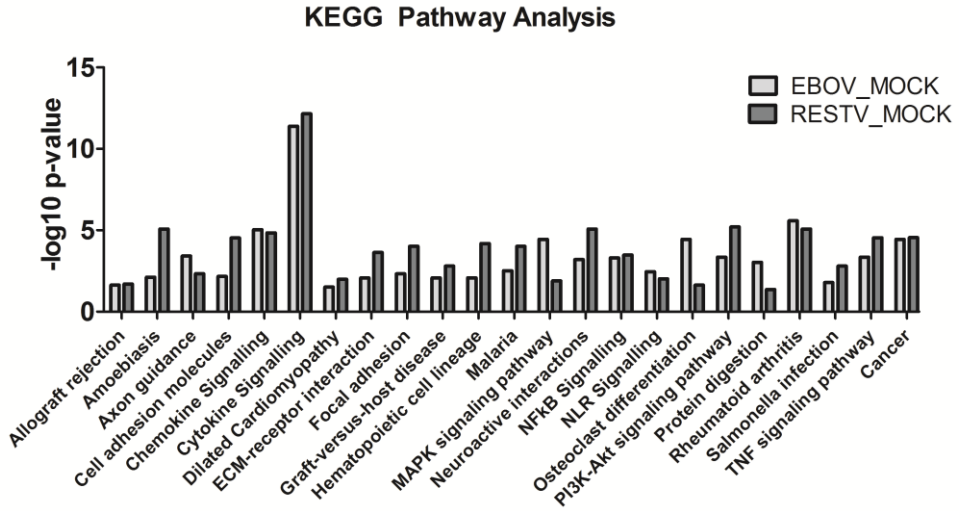


Figure 6.7: KEGG pathway analysis performed using the online KEGG Mapper toolkit applied to protein abundance data from either EBOV or RESTV infected THP-1 cells. Only proteins with significance abundance measurements above mock infection were selected, and these were then filtered for those proteins with corresponding transcript abundance data. The significance of protein alignment to a given functional KEGG pathway was demonstrated by the corresponding -log10 p-value. The greater this -log10 p-value the greater the significance. Data was separated into either EBOV_MOCK or RESTV_MOCK and indicated by colour shown in the legend. The name for each KEGG pathway identified is shown on the x-axis.

6.3.4 Proteins involved in viral infection show significantly different levels of protein and transcript production.

Proteins with significant abundance changes were contrasted filtered for those with corresponding transcriptomic data. Functional network analysis using STRING 10 identified a sub-network of host proteins involved in virus infection. These 11 proteins also showed a dichotomy in the level of protein abundance identified in comparison with transcript abundance. The significance of this was questionable, as it was known that transcript abundance measurements do not always correlate with protein abundance. However the analysis provided useful targets for western blot validation of proteomics analyses. The results of this analysis were shown in Table 6.2

6.3.5 Inflammatory and immune response proteins change in abundance during viral infection.

To confirm the levels of proteins identified in proteomic analysis, a western blot targeting specific host proteins of interest in whole cell lysates was performed. Included in this analysis were VDAC and VIM which were identified in a sub-network of proteins involved in viral infection. As the transcriptional regulatory complex NFκB was also suggested to be significant in proteomic and transcriptomic analysis, evidence of complex activation was sought. Antibodies specific for phosphorylated p65 and IkkB were used to indicate the presence of NFκB complex activation. Interestingly the level of IkkB did not discernibly change, and the level of phosphorylation of p65 appeared reduced in infected cells compared with non-infected. The results of this western blot are shown in Figure 6.8. Phosphorylation of p65 is due to cytokine mediated activation of PKA, which phosphorylates p65, inducing a conformational change, distinctly changing the repertoire of transcriptional promotion from unphosphorylated p65, indicating native p65 was abundant.

Gene Name	Symbol	RESTV RNASeq	RESTV Proteomics	EBOV RNASeq	EBOV Proteomics	Function
Rnase L Inhibitor	ABCE1	0.18	1.04	0.81	-0.09	ATPase Activity
Charged Multivesicular Protein 48	CHIMP4 8	0.21	0.27	0.49	-0.11	Endocytosis
High Mobility Group AT-Hook1	HMGA1	-0.09	1.23	1.57	0.24	Transcription Regulation
HSP70	HSPA8	0.16	1.73	1.78	0.28	Chaperone
HSP60	HSPD1	-0.24	0.99	1.06	0.02	Chaperone
Integrin Subunit Alpha 5	ITGA5	0.31	0.71	0.98	-0.02	Cell Adhesion
Karyopherin Alpha Subunit 3	KPNA3	0.14	0.45	0.27	0.19	Nuclear Transport
RNA Guanine-7 Methyltransferase	RNMT	0.26	0.52	0.51	0.06	Transcription Regulation
Ribosomal Protein L34	RPL34	0.21	-0.08	-0.68	0.50	Ribosomal Function
Voltage Dependent Anion Channel 1	VDAC1	-0.35	0.74	0.63	0.21	Mitochondrial Function
Vimentin	VIM	0.63	1.70	1.74	0.44	Cytoskeletal Function

Table 6.2: Table showing a sub-network of proteins identified using STRING 10 related to the antiviral response. These 11 proteins were compared with transcript abundance measurements, shown is log2 fold change values for both transcripts and protein level abundance data. There is a lack of correlation in most cases. This may be due to modulation of protein production at the post-transcriptional level, or may be a temporal issue. Listed were protein/gene names under the Gene Name heading, corresponding gene symbols are shown, alongside results for RNASeq and proteomic analysis. A description of function is also provided, alongside the GO term associated with this sub-network.

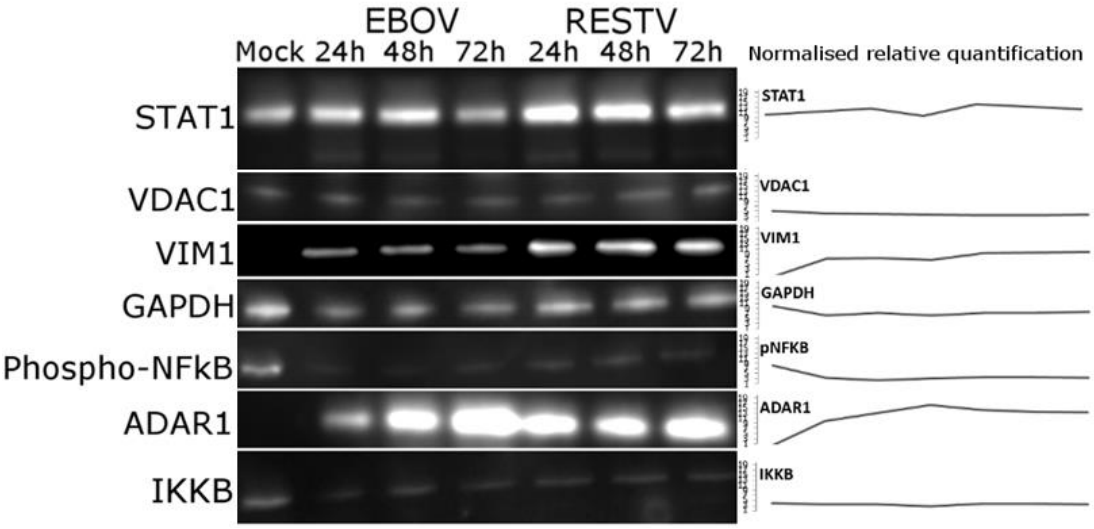


Figure 6.8: Western blot analysis of 7 proteins identified through proteomic analysis. Samples were collected and lysed in Laemmli buffer supplemented with 10% SDS. Samples were arranged on the gel as follows: mock infected, EBOV infected cells at 24h, 48h and 72h, then RESTV infected cells at 24h, 48h and 72h. Shown are results for STAT1 VIM1, ADAR1, IKKB and Phosphorylated-p65 (NFκB). Housekeeping protein GAPDH is used as a loading control, and is shown as a representative example. Each test included GAPDH as a loading control. VDAC1 was also considered a housekeeping protein, but was identified as an associated protein in STRING 10 analysis, alongside antiviral response proteins. The intensity of bands correlates roughly with the abundance of protein loaded onto the gel prior to blotting. Transfer of blot was controlled with PureLink pre-stained standard, and sizes on western blot analysis controlled by MagicMarker fluorescent standard.

6.3.6 Cytokine and chemokine secretion correlates with transcriptomic predictions of activity.

Prominent changes in the abundance of transcripts encoding cytokines and chemokines in infected THP-1 cells were identified through transcriptomics but were not detected in proteomic analysis. The reason for this could be multifactorial but was most logically because such proteins are readily exported from the cell during activation. To compensate for this a multi-analyte immunoassay screen for important regulatory cytokines and chemokines including IL1B, IL6, IL12 and TNF was performed on supernatant collected from infected cells at 72h post infection. Samples assayed included those collected from EBOV and RESTV infected cells, commercially produced control cells treated with LPS and negative control THP-1 mock infected cell supernatant.

The analysis shows that there was detectable increase in production of TNF, IL1B, IL8 (CXCL8) and RANTES. Production of RANTES (CCR5) was of particular note, being far in excess of that detected in the commercially provided control material. Curiously, there was only limited secretion of interferon-α (IFNA), in contrast to detected transcript abundance measurements. A graph of these ELISA results were shown in Figure 6.9.

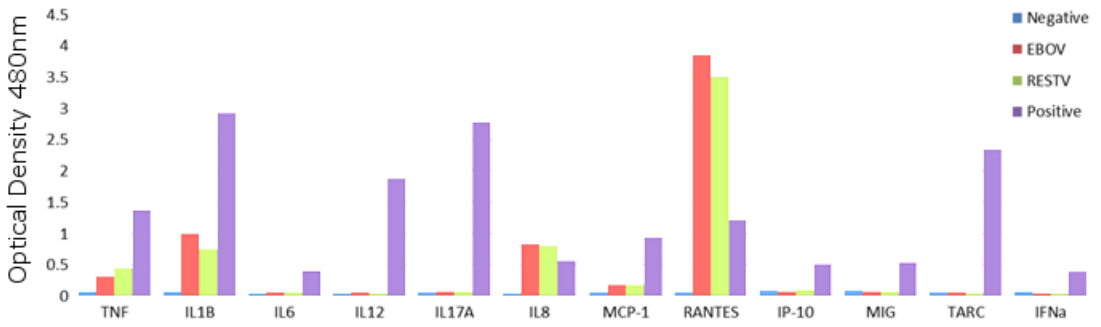


Figure 6.9: Results of a multi-analyte ELISA analysis of supernatant collected from EBOV and RESTV infected THP-1 derived macrophages at 72h post infection. Shown are 12 proteins identified in solution by capture ELIA. These are arranged on the x-axis as TNF, IL1B, IL6, IL12, IL17A, IL8, MCP-1, RANTES, IP-10, MIG, TARC and IFNα. The abundance of the protein identified in supernatant correlates with the optical density measured by Spectrophotometer at 480nm. Negative control (Mock infected) is shown alongside supernatants from EBOV infected cells, RESTV infected cells and positive control (from LPS stimulated U937, commercial, in-kit analyte).

6.4 Identification of host factors.

6.4.1 Use of differential transcriptomic analysis to select targets for inhibition

Combined transcriptomic and proteomic analysis of EBOV and RESTV infected THP-1 cells was reviewed alongside data acquired from EBOV and RESTV infected A549 cells. This review was aimed at defining high priority targets for functional analysis. Initially specific genes and proteins were selected but later disqualified due to limited data on downstream network influences. Instead, upstream regulators common to transcriptomic and proteomic analysis in both A549 and THP-1 cells were selected.

NFκB complex, TNFα, and the upstream transcription factor SP1 were identified as having a common role in all data, and well informed pathway analysis demonstrated wide ranging network influences on signalling pathways. SP1 is a transcriptional regulator with wide ranging roles in cell biology, the essential role for this protein made it unlikely to have direct influence over viral lifecycle, but downstream proteins under its regulatory control may have a role to play. NFκB has significant control of cell survival, inflammation and immune response. Published works have repeatedly highlighted the importance of NFκB in the lifecycle of clinically important viruses. The role of NFκB in *ebolavirus* lifecycle has been suggested previously [221].

6.4.2 Selection of compounds for inhibition of SP1 and NFκB

Inhibitors with targeting specificity for the NFκB complex or SP1 were identified using the inhibitor database. Withaferin A has broad specificity on multiple cell networks under the influence of SP1, evidence of direct effects on SP1 activity suggested reasonable levels of specificity for this transcriptional regulator. Withaferin A is a compound derived from *Withania somnifera*, commonly known as Indian Winter Cherry. The compound has routes in traditional ayurvedic medicine and is licenced for traditional medicine use in the UK and Europe. The compound has proven efficacy at reducing inflammatory response and is immunomodulatory, a review of the multifunctional activity is available from Vanden Berghe et al 2012. Withaferin A was therefore chosen to as an inhibitor of SP1 to measure the influence of this regulator on *ebolavirus* lifecycle. A synthetic chemical inhibitor SM7368 also designated JSH-23 is designed to inhibit the nuclear translocation of the NFκB complex, thereby acting to reduce the inflammatory and pro-survival effects of NFκB by reducing TNFα and TLR4 induced NFκB transcriptional activity. SM7368 is not licenced for use in humans, but drug specificity made it an attractive option for inhibition testing. Schematic diagrams of the chemical structure of these compounds is shown in images A and B of Figure 6.11.

6.4.3 Cytotoxicity assay for inhibitors in HEK-293T

Cytotoxicity of SM7368 and Withaferin A were compared with DMSO and PBS to identify negative cell effects. This analysis illustrates the detrimental effects of both compounds on THP-1 cell viability. HEK-293 cells however seem to retain viability in the presence of high concentrations of drug. Thus HEK-293 cells were used for inhibition assays, as these cells were also readily transfectable with mini-replicon components (figure 6.14)

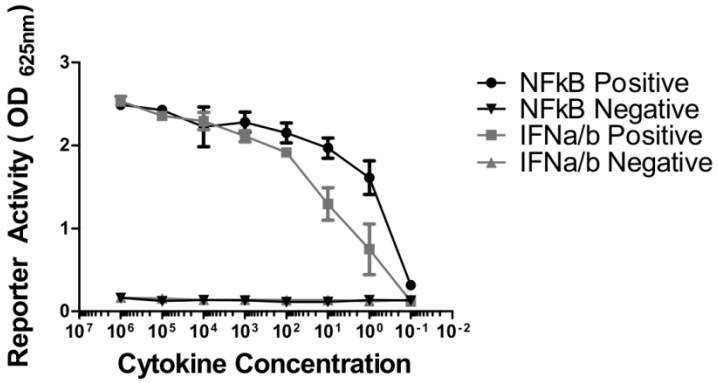
6.4.4 Selection of HEK-293 cells for inhibition testing

HEK-293 cells were used in inhibition assays as mini-replicon transfection was possible in these cells but did not yield effective luciferase measurement in THP-1 cells. The reasons for this may be that THP-1 macrophages were refractory to transfection with DNA/Plasmids due to the presence of high concentrations of DNases and exogenous nucleic acid sensing pathways (such as STING) contained within macrophage-like cells. Additionally, cytotoxicity assay reports indicated that HEK-293 cells were more resistant to chemical treatment with SM7368 and Withaferin A compared with THP-1 macrophage-like cells. The results of cytotoxicity testing for HEK-293 and THP-1 derived macrophage exposed to different concentrations of WFA and SM7368 is shown in Figure 6.12.

6.4.5 Validation of assay stimulation reagents

To verify that test reagents used in the assay were eliciting measurable stimulation of NFκB or IFNα/β, reporter cell lines were used. HEK-Blue cells (HEK-293T derived cell line) transfected with either the IFI53 promoter driven by stimulation with IFNα/β or promoters under the power of NFκB and AP-1 transcription factors, which are activated in response to stimulation with TNF. Additionally, cells transfected with luciferase under the power of NFκB and AP-1 were treated with an inhibitor (SM7368) prior to treatment with stimulating reagents. The results of this experiment are shown in figure 6.10, illustrating that treatment with TNF or IFNα/β results in a measurable increase in luciferase activity, and treatment with SM7368 reduces the level of detectable luciferase activity in cells where production is under the power of NFκB.

A.



B.

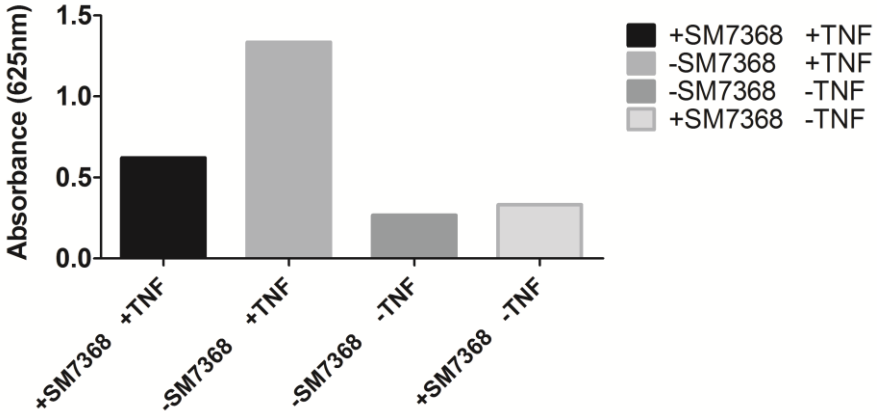
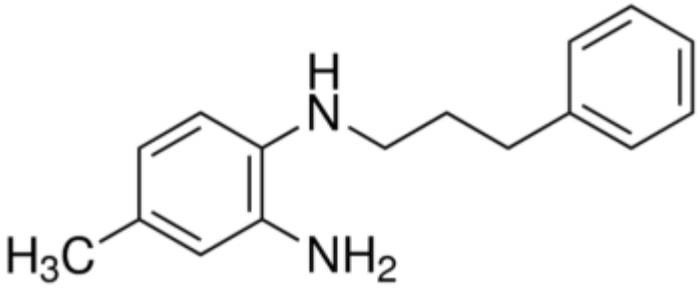


Figure 6.10: Results of experimental treatment of SEAP-Reporter producing cell lines with stimulating reagents. (A) Shows treatment of both NFκB and IFN reporter cell lines treated with concentrations of stimulating reagent (either TNF or IFNa) a concentrations ranging from 10e+7 pg/ml to 10e-2 pg/ml. Results shown are an average of 3 experiments. (B) shows the results of experiment testing the efficacy of the SM7368 NFκB inhibitor on reporter cell line SEAP activity. Results shown represent a single experiment only.

A.



B.

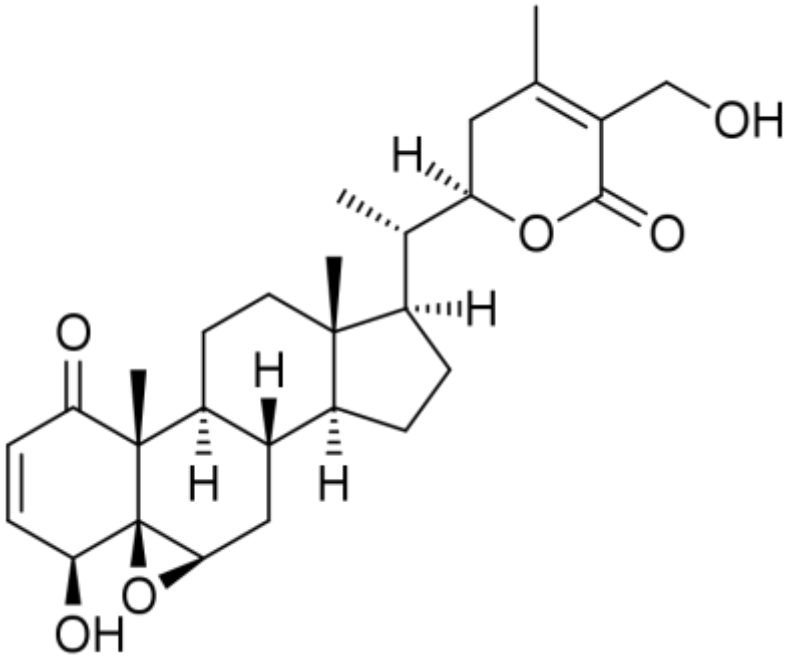


Figure 6.11: Schematics of chemical structure of SM7368 and Withaferin A (WFA). (A) The chemical structure of SM7368 shows a low complexity compound with a molecular weight of 328.69. SM7368 is capable of entering cells and inhibiting TNFα mediated MMP9 production, and a potent inhibitor of TNFα induced NFκB activation. (B) The chemical structure of Withaferin A (WFA), a natural compound extracted from plant material (Indian Winter Berry), with traditional herbal use as an anti-inflammatory in ayurvedic medicine. The compound is structurally complex and the exact mechanisms or the extent of interactions were unknown. WFA is confirmed to interact with the broadly-active transcriptional factor SP1.

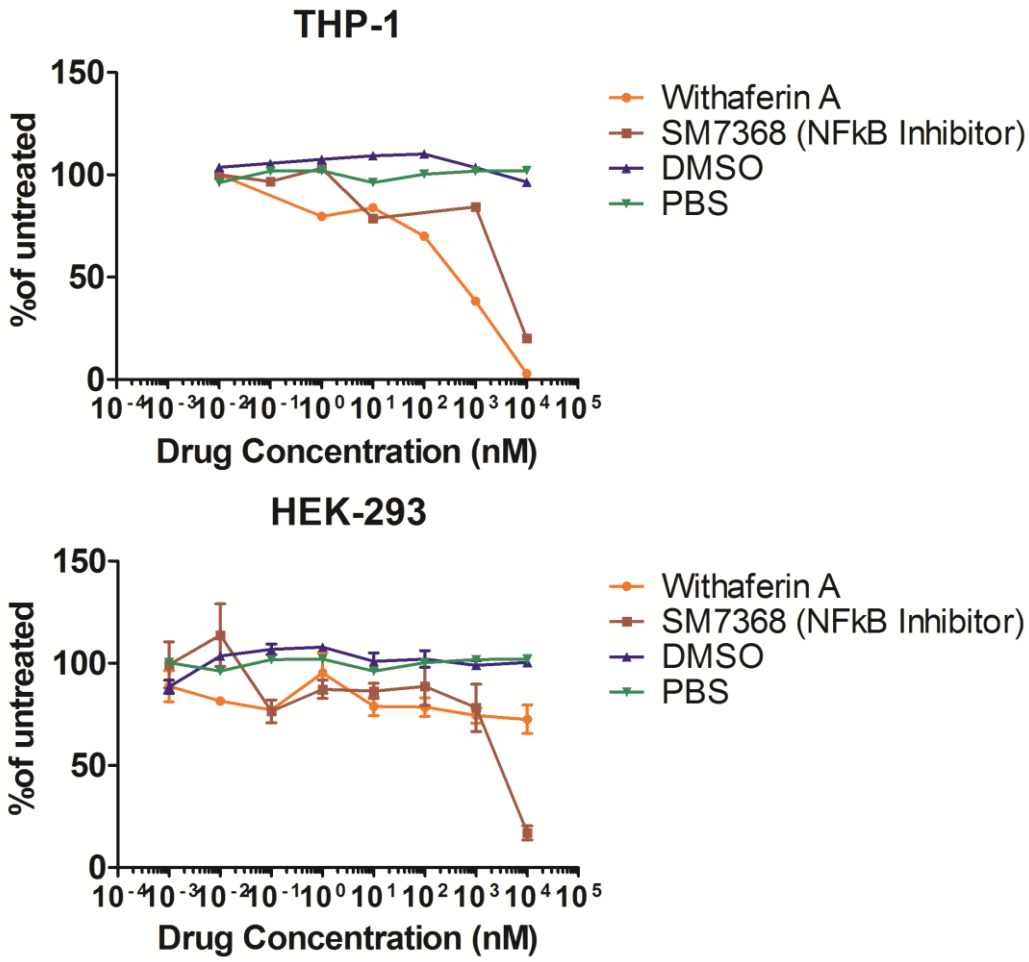


Figure 6.12: Cytotoxicity assays performed on HEK-293 and THP-1 derived macrophages treated with increasing doses of Withaferin A, SM7368, DMSO or PBS. The y-axis indicated the percentage of untreated cells (no treatment) which remained viable in the cytotoxicity assay. This analysis reveals that PBS and DMSO show no discernible toxicity, as expected. However there is toxicity of WFA and SM7368 in a dose dependent manner. In excess of 1µM concentration there is severe decline in cell viability observed in THP-1 cells treated with either compound. However HEK-293 cells remain above ~80% viability when treated up to 10µM SM7368.The scale of nM concentration is shown on the y-axis.

6.4.6 Luciferase activity measurements of mini-replicon system at different concentrations of inhibitor in HEK-293 cells

Mini-replicon activity was measured in HEK-293 cells cultured in the presence of high concentrations of inhibitors. Cells were treated with 1µM Withaferin A defined as the highest concentration of inhibitor shown to be usable while retaining >80% cell viability. Withaferin A showed a slight increase in EBOV and RESTV mini-replicon activity as measured by luminescence assay. The levels of increase were analysed by ANOVA and shown to be insignificant (Data not shown).

Next HEK-293T cells were exposed to either 1µM of SM7368 NFκB inhibitory compound, or 100nM TNFα, or treated only with DMSO (Neg). This showed a significant decrease in mini-replicon activity when treated with SM7368 compared with negative control. TNFα treatment showed a slight decrease compared with negative control, but was not significant. The results of this analysis are shown in Figure 6.13.

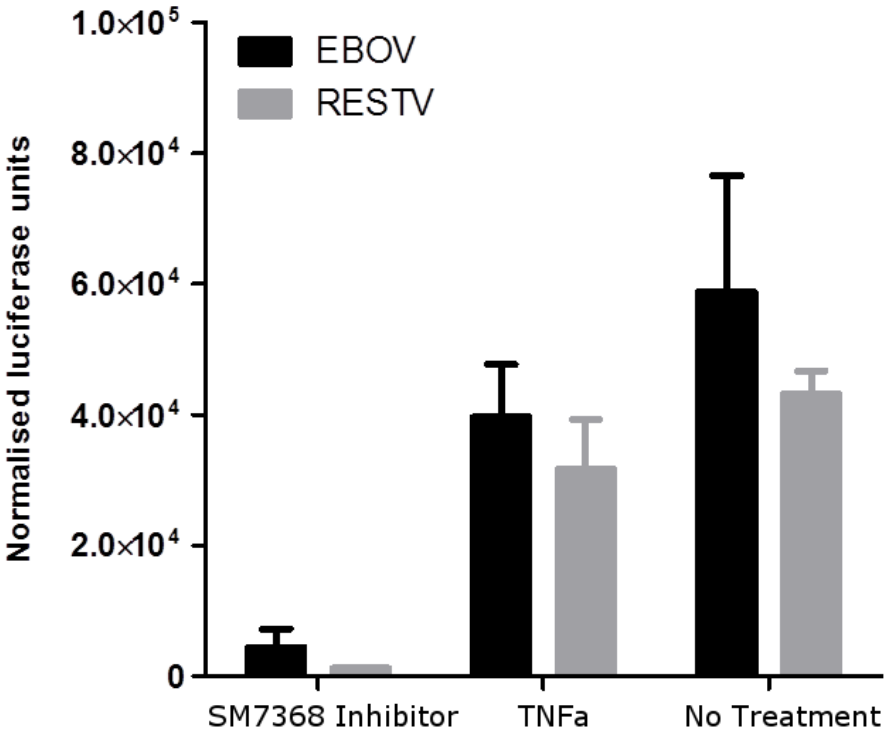


Figure 6.13: Results of a luciferase measurement assay performed on cells transfected with the mini-replicon models of EBOV (Makona) or RESTV, and treated with SM7368, TNFα or untreated (DMSO vehicle only) in triplicate transfections (with triplicate transfected wells). The y-axis indicated the luciferase units normalised to control transfected cell measurements (transfections performed with the absence of the L-polymerase encoding plasmid). The amount of luciferase detected was proportional to the replicative activity of the mini-replicon system in HEK-293 cells. This indicated that the amount of luciferase activity was reduced by treatment of cells with SM7368, slightly reduced by treatment with TNF, and greatest when untreated. This analysis was based on results of separate transfection experiments (n=3), error bars are based on the standard deviation of results. No significance testing was performed due to the relativistic nature of results, and the small number of replicates.

6.5 Conclusions

The focus in this chapter was the study of the effects of the EBOV Makona variant on a model of macrophages using proteomics. A secondary objective of this chapter was to ascertain whether these changes induced in the host response to EBOV Makona were specific to this virus, or shared with the non-pathogenic RESTV. The second part of this chapter aimed to address whether identified host factors were significant, by using inhibitors to evaluate effects on virus replication.

The proteomic analysis performed in this chapter builds on the transcriptomic analysis performed in chapter 5. Combining data from transcriptomic and proteomic analysis made found that there was significant evidence of increasing abundance in inflammation associated proteins and transcripts. Pathway and network analysis repeatedly revealed in chapters 3, 4,5 and 6 that NFκB may have a significant association with infection with both pathogenic and non-pathogenic *ebolaviruses*. To determine whether NFκB activation was important for effective virus replication a specific inhibitor of the P65 subunit of NFκB, and the anti-inflammatory Withaferin A were used to interrupt inflammatory processes. The effects of these compounds were measured by using a mini-replicon model system of EBOV and RESTV, and the results of this analysis indicated that NFκB activity may be important in both EBOV and RESTV lifecycle.

7 Thesis Discussion

Prior to the research conducted in this thesis the standard paradigm for the interaction of EBOV with the host was that the virus was very efficient as shutting down the host cell antiviral response and modulating the host organism's immune response by antagonising interferon signalling pathways, stimulating pro-inflammatory immune cells, and increasing cell activity through induction of pro-inflammatory signalling pathways (24, 63, 78, 83, 102, 104, 114, 117, 135, 154, 155, 158, 163, 190, 200, 202, 203, 206, 218, 220-224). The scale of the outbreak in West Africa allowed hitherto unknown aspects of virus biology to come to the forefront, particularly the idea of virus persistence (225). Although a related virus had been identified in the semen of an accidently infected laboratory worker many years before (12). One of the other major shifts in research between intervening outbreaks was the development of high resolution approaches such as RNA sequencing and quantitative proteomics. RNASeq data utilised in these studies have been carefully interpreted, the technology allows exploration of molecular processes in exquisite detail, but may suffer from technical bias; broad claims about the data have been avoided in favour of a targeted analysis. The research conducted in this thesis was complemented by deployment to the outbreak in West Africa to aid in the diagnostic effort and also the use of diagnostic left over samples for research purposes. Rather than relying on purely cell culture studies, the samples from human infection provided a useful comparator. Undoubtedly, cell culture remains a tractable system for studying EBOV biology, particularly replication and transcription. The data from thesis highlights the need to consider cell culture data alongside clinical data to study all aspects of the host response.

7.1 Longitudinal analysis of the Makona variant of EBOV in a patient with EVD

The detailed analysis of a clinical case of EVD has highlighted antiviral signalling as having a close relationship with viral load in EVD patients. Though host transcription may have been affected by treatments given, the patient was not treated with interferon, and thus it can be determined that interferon response was host derived. Canonical pathway analysis showed significant alignment of genes to inflammatory and antiviral pathways. A strong interferon response was observed early in infection, similar to that previously described in patients and animals (117, 132). This EVD case was rare as the patient suffered clinical signs of a likely fatal infection, yet survived (likely due to effective life support). From cT values alone (peaking at <cT 15), had this patient been treated in an outbreak setting with minimal life-support, the patient would not have survived. Though other studies have suggested that time to death in Makona infection may be extended in humans and macaques (200, 226, 227), the clinical descriptions of this EVD case appeared very similar to previous outbreaks (117, 190, 207, 228, 229). Sequencing of viral genomes across the outbreak allowed epidemiologically important discoveries, but did not suggest a fast rate of adaptation or mutation. EBOV appeared to be a stable organism, even in outbreaks amongst non-reservoir species. Descriptions of clinical disease, and cellular transcriptomic changes in Makona infection were very similar to Ecran infection in this a study of both viruses in cell culture. Therefore considering this combination of *in vitro* and clinical evidence, these viruses should be considered similar.

7.2 The significance of NFκB signalling in the lifecycle of the Makona variant of Ebola virus.

High resolution transcriptomics of the host inflammatory response to Makona correlated with previous attempts to understand the host response to other EBOV isolates (Chapters 3, 5, 6).

IL1B, IL6, IL8 and IL15 hypersecretion has previously been suggested to correlate with lethal clinical outcome (117). IL1B and IL15 showed a significant increase in Ecran compared with Makona (Chapter 3). IL6 and CXCL8 (IL8) showed similar patterns of transcript abundance in both EBOV variants and RESTV (Chapter 3, 4). Secretion of these cytokines and chemokines were associated with stimulation of cells by TNF, and can also be associated with induction via TLR4 mediated activation of the NFκB complex. Differences in the levels of these cytokines may suggest a slight difference between these viruses in TNF or NFκB activation (Chapters 3, 4). Metalloproteases associated with TNF/NFκB also appeared raised in both Ecran and Makona (Chapter 3), which have been previously associated with poor clinical course in the mouse model (143).

These initial studies of EBOV and RESTV viruses in A549 cells highlighted the TLR4, TNF and NFκB mediated signalling pathways as potentially important in *ebolavirus* infections (Chapters 3 and 4). THP-1 cells, a cell type to model inflammatory response, were selected to analyse the TLR4 mediated inflammatory response using a biologically relevant cell type (Chapter 5 and 6). THP-1 cells displayed typical phenotypic changes when differentiated into macrophages by previously described protocols (178-180, 211, 212). Similar patterns of expression were identified which suggested the THP-1 cells expressed known macrophage differentiation markers (180, 215, 216). EBOV

and RESTV infected THP-1 cells at similar efficiency, and replicated at similar rates in THP-1, A549 and HEPG2 cells infected at an MOI of 5 (Chapter 5). Transcripts encoding pro-inflammatory cytokines were highly expressed, and many changed significantly in abundance (Chapter 5, 6). This high activity is characteristic of macrophages when activated. The observed increase in TNF and IL1β described in THP-1 cells was notable as these regulators appeared slightly elevated in EBOV infection of A549 cells compared the Ecran variant of EBOV (230), as described in Chapter 3. IL6 has previously been described as upregulated in patients and animals infected with EBOV (117, 124). This strong pro-inflammatory cytokine response was reflected in the analysis of upstream regulators, and was raised in all cellular infections with Makona. IPA analysis highlighted NFκB, TNF, IL1β, TGFβ and LPS (TLR4) as top regulators of significantly changing transcripts (Chapters 3 and 5). The roles of NFκB and TLR signalling have been highlighted in other studies and were believed to be due to the interaction of viral glycoprotein with TLR4 (135, 161-164). Single-site analysis using oPOSSUM revealed that SP1 and NFκB were major regulators of transcript abundance changes induced differentially by EBOV and RESTV compared with mock infection. Proteomic analysis further illustrated the role of NFκB and SP1 in EBOV and RESTV infection of THP-1 derived macrophages by demonstrating significant abundance changes in proteins regulated by these host factors. Combined proteomics and transcriptomics showed a correlation between numerous proteins and transcripts under the regulatory control of NFκB which had changed significantly in abundance (Chapters 5 and 6). Inhibitor experiments using SM7368 demonstrated that inhibition of NFκB resulted in a reduction of EBOV and RESTV replication.

Similar effects have been observed in influenza A virus infection, requiring NFκB pro-inflammatory activity to effectively replicate (131, 231, 232). Indeed, concurrent research into the host response to EBOV and RESTV demonstrated a key variance in the relationship of these viruses to NFκB activation (220). This study deviates from the work described in chapter 5 in several ways. Rather than THP-1 derived macrophages, primary cells have been used. Most importantly read depths of ~10 million reads per sample were reported, in contrast to this study, in which average accumulated read depth from next-generation sequencing was in excess of 60 million per sample. This depth may have resulted in estimated increased in relative abundance, while lower read depths may not have highlighted these transcripts as significant. Additionally, THP-1 cells are more representative of inflammatory macrophages distinct from resting ex vivo macrophages which have a separate cell lineage. The origin of this variance in inflammatory activation appeared to be due to differences in the activation of TLR4 by GP, which was also predicted to contribute to differences in pathogenicity *in silico* (105). Careful comparison of the methodologies demonstrated that a key difference was the way in which virus inoculum was prepared. While virus stocks were diluted in culture media for inoculation in all cell culture infections described in this thesis (Chapters 3, 4, 5, 6), virus used in the concurrent study was prepared by ultra-centrifugation, suggesting the possibility that an as yet undefined characteristic of infected cell products are influencing background changes in inflammatory state. RESTV attenuation in human infection may in fact be due to reduced pro-inflammatory activation. A model may be proposed illustrating the role of NFκB in *ebolavirus* lifecycle, shown in Figure 7.1. Additionally, increased activity of EIF2α signalling was highlighted in Chapters 3, 4 and 5.

Other transcripts involved in the stress granule response have been observed to be activated by the Makona variant of EBOV. Sensing of double stranded RNA by PKR results in EIF2α phosphorylation, resulting in inefficient processing of viral proteins. The role of the cellular stress response in EBOV infection has yet to be explored in significant detail, but may be a useful area for further research. Transcriptomic and proteomic analysis have identified CD40/CD40L, TNF, IL1R, TLR4 and growth factor associated transcripts and proteins which may be contributing to a potent NFκB response in an additive manner. In this proposed model a overstimulation of the NFκB response results in hypersecretion of inflammatory cytokines by affected cells, and contributes to the prevention of cellular apoptosis which has been observed in previous studies (134) by promoting cell survival. Hypersecretion of pro-inflammatory cytokines have deleterious effects on cells of the adaptive immune response, and may explain the bystander cell death observed affecting lymphocyte populations, which resist infection with EBOV but were susceptible to induced cell death (109, 117, 119). Prediction of which downstream host factors may be contributing to the enhanced growth of *ebolaviruses* upon NFκB activation was difficult. The pro-inflammatory p65 component of the complex was recently shown to be a major contributory element of the observed requirement for NFκB activation to permit influenza A infection alongside c-Rel co-factor (233). The high degree of pro-inflammatory signalling downstream of the NFκB complex highlights p65 regulated genes as interesting targets for future study. Comparative analysis of the 1918 variant of influenza A virus suggested that this variant of IAV showed increased NFκB induction, and thus the strength of NFκB induction may correlate with the pathogenicity of the infecting virus (174, 234, 235).

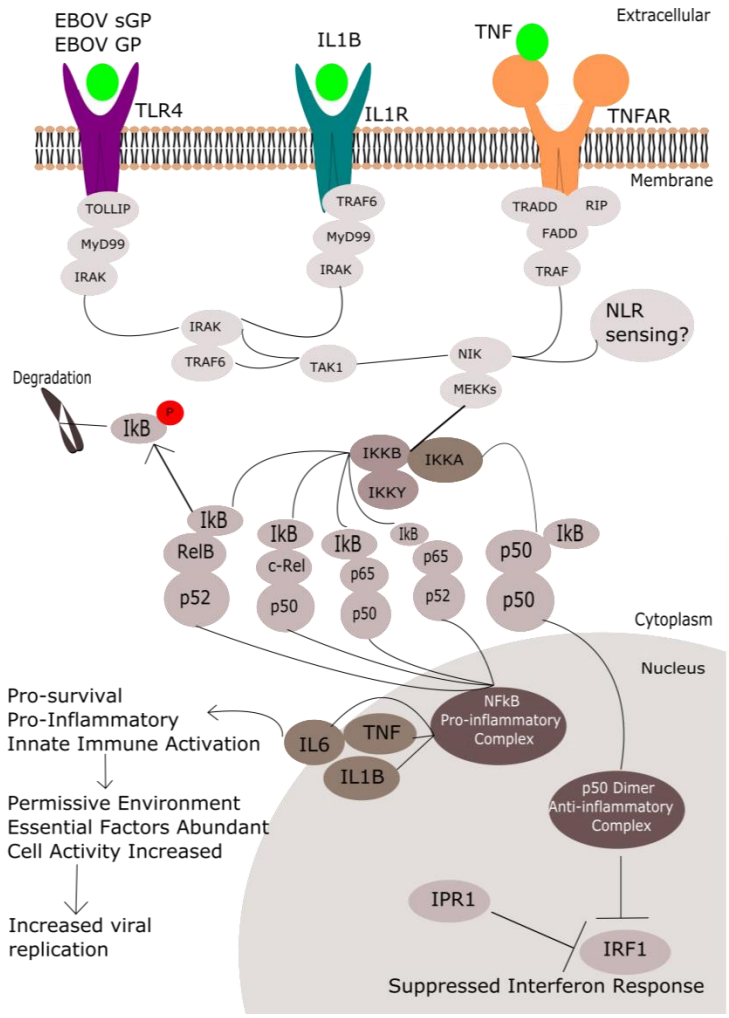


Figure 7.1: Schematic diagram of hypothesised model of the relationship between EBOV infection and NFκB activation. Activation occurs via stimulation of multiple receptors including TLR4 (by EBOV sGP/GP), IL1R by IL1β and TNFAR by TNFα. Co-factors activate in a cascade of signalling, activating MEKKs, activating IkB kinase IKKB resulting in the phosphorylation of IkB inhibitory proteins, and the release of NFκB complexes to translocate into the nucleus. Heterodimeric complexes composed of p65, p50, 052, c-Rel, RelB are pro-inflammatory, homodimers may also form from p50-p50. P50 homodimers are anti-inflammatory and act as a negative feedback system controlling NFκB pro-inflammatory activation. P50 homodimers suppress the activity of IRF1, a key type 1 interferon regulator, further suppression by IPR1 and the EBOV proteins VP24, VP35 render the type 1 IFN response ineffective. Pro-inflammatory proteins are produced, NFκB heterodimers result in pro-survival signalling, preventing apoptosis. Increased cellular activity and upregulation of as yet unidentified host factors results in increased viral replication rates, and improved permissiveness for viral infection.

7.3 The significance of interferon in *ebolavirus* infection.

EBOV is known to produce potent interferon antagonist proteins (101). A549 cells were selected for their potent interferon response (192, 235, 236), in order to better assess the impact of Makona and Ecran on interferon signalling in a competent cell line. Interferon stimulated transcripts seemed to have similar transcript abundance profiles for both Makona and Ecran. This correlates with other data suggesting the main interferon antagonist proteins of Makona were equally effective at antagonising the interferon response (237). The interferon response in cells infected with EBOV *in vitro* was completely different to that observed in human and animal hosts, this conforms to previous observations [118, 130]. Indeed, viral load correlates with increased abundance of transcripts encoding antiviral response genes, as shown in Chpt 3, figure 3.5 (p138). The reasons for this were not fully explored; it was likely that interferon signalling was suppressed in infected cells, while non-infected bystander cells react to other intra-cellular signalling events in response to viral infection, only detectable in complex systems such as *in vivo*. Previous work has highlighted potential differences in the ability of RESTV to influence interferon signalling events compared with EBOV (160). Indeed, it RESTV interferon signalling antagonistic proteins VP24 and VP35 possess similar function and structural characteristics; here it is demonstrated that RESTV and EBOV induced similar interferon signalling in human and animal cell lines. This indicated that the Makona variant of EBOV does not possess interferon signalling antagonistic capabilities in excess of what has been expected from other variants, and that, in any case, interferon does not appear to determine the outcome of infection; RESTV was has yet to cause discernible disease in humans yet patterns of interferon signalling are similar to EBOV.

Other haemorrhagic fever causing viruses including Lassa fever virus [13, 239] and Crimean-Congo haemorrhagic fever virus [240,241] are capable of inhibiting the antiviral response of human cells without perturbing the inflammatory processes which contribute to the disease they cause [242].

Investigation of the antiviral response in infected THP-1 derived macrophages by qRT-PCR and RNASeq revealed a stronger inflammatory response to EBOV than RESTV. However interferon signalling activity was low compared with that observed in A549 and HEPG2 cells, and interferon associated transcript abundance were similar in both EBOV and RESTV infection. Therefore THP-1 cells have a weak interferon response, though previous studies have highlighted the ability of THP-1 cells to mount a strong antiviral response to other viruses (213, 214). This may suggest the interferon response was suppressed by both EBOV and RESTV from the outset of infection, in concordance with previous studies (203). Multi-analyte ELISA results add evidence to this claim; measurement of the interferon response showed no activity at the protein level; as no significant interferon-α production was detectable after infection of these cells. Caution is advised however; as no published studies have confirmed that THP-1 cells were capable of secreting type 1 and 2 interferons, though transcriptomic and qRT-PCR data suggested they may. In conclusion the interferon response does not appear to be a determining factor in the ability of *ebolavirus*es to infection host cells; human non-pathogenic and human pathogenic *ebolaviruses* appear capable of inhibiting the interferon response with similar efficiency.

7.4 Additional outcomes

7.4.1 Unusual patterns of virus transcription

The abundance of viral mRNA was compared between the variants to confirm qRT-PCR measurements of viral genome abundance and examine whether the synthesis of Makona and Ecran mRNAs differed in pattern or abundance. This comparison was performed again during infection of A549 and THP-1 cells with EBOV or RESTV, to control for viral abundance. For members of the *Mononegavirales*, to which EBOV and RESTV belong, synthesis of mRNAs has been postulated to reflect a transcription gradient (60), where genes transcribed from the genomic 3' end were more abundant than genes transcribed at the genomic 5' end. However, recent work has suggested the transcription pattern for viral mRNA transcribed by Hendra virus and HRSV (151, 238), both members of the *Mononegavirales*, does not conform to this expected pattern. Mapping of EBOV genes was conducted as previously described for HRSV (238), EBOV in a guinea pig model (104) and patients (193), and described in the methods chapter of this thesis. The abundance of the NP mRNA was greater than that of L mRNA for both viruses; however the abundance of the viral mRNAs transcribed between these genes was not linear, and did not conform to the expected gradient pattern. Indeed there was a substantial difference in transcription of the L polymerase compared with the rest of the genome in studies described in this thesis. The existing model of *ebolavirus* transcription should be re-evaluated in light of these observations.

7.4.2 Species affinity of RESTV

It has long been known that RESTV was capable of causing disease in a species dependent manner. Infection of NHP with RESTV results in disease, all be it slightly milder than would be observed for EBOV infection (27).

In this study RESTV displayed accelerated replication in porcine cell lines compared with EBOV. This was not due to a reduced interferon response to RESTV in these cell lines compared with EBOV (Chapter 4). qRT-PCR analysis of 84 antiviral associated transcripts showed greater abundance of those encoding inflammatory proteins, suggesting a correlation between increasing inflammatory response and virus replication. This further evidenced the necessity of NFκB activity in *ebolavirus* lifecycle, in addition to data from Chpt 4,5 and 6. The observation that RESTV may be more suited to replication in porcine derived cell lines was intriguing as RESTV has been found to cause epizootics in domesticated pigs (25, 26, 29). The data in Chpt 4, 5 and 6 demonstrate that RESTV is fully capable of infecting human cells and initiating a pronounced immunopathology *in vitro*, while RESTV has never been reported as a cause of EVD in humans, this *in vitro* evidence represents an assessment of potential risk, and RESTV should not be considered any less dangerous to humans than other *ebolaviruses*.

7.4.3 Development of a model system

The requirement arose to develop a system which would allow measurement of EBOV and RESTV without relying on access to high containment facilities and expertise. It was desirable to implement such a model to allow testing of inhibitors, drugs and compounds which may influence EBOV lifecycle by targeting host protein activity. The monocistronic mini-genome previously published in the literature was adapted (72, 100), and created using the sequence data from isolates of the Makona variant of EBOV and the available isolate of RESTV.

7.4.4 Identification of therapeutic compounds

The identification of NFκB allowed the verification that inhibition of activity was able to interfere with EBOV and RESTV lifecycle. This indicated

a broad spectrum of effectiveness. Though SM7368 was not licenced for treatment use other alternative drugs were available with activity against NFκB may be suitable for investigation as a broadly efficacious therapeutic. This was especially true as NFκB signalling appeared to underpin most signalling events during cellular infection with EBOV (Chapters 3-5), and knowledge of clinical disease may make anti-inflammatory therapeutics relevant (117, 129). This work contributed to a Wellcome Trust funded drug discovery project, with contributions shown under associated publications.

7.5 Final Conclusions

The core aim of this project was to characterise the host response to the recently discovered Makona variant of EBOV, comparing the profiled changes in the host response with those induced other variants of EBOV and other *ebolaviruses* of differing pathogenicity, helping to identify transcriptomic and proteomic signatures of infection with the emerging Makona variant. The analyses described in this thesis illustrated the surprising similarity in the transcriptional profiles found in cells infected with the EBOV viruses tested. Likewise many similarities in the host response were observed in RESTV infection, a pathogenically dissimilar *ebolavirus*. It was shown that the interferon response is not likely to be a major determinant of successful *ebolavirus* infection. The main determinants of cell permissiveness were shared by both EBOV and RESTV, and that the host factors related to this were likely to be under the regulatory control of the NFκB complex. Maintaining a potent inflammatory reaction may be critical in *ebolavirus* lifecycle, and correlates with the early observations that EBOV Makona mutation during the West African EVD outbreak did not support the theory that the virus was becoming less pathogenic.

It was hypothesised that the host response to the Makona variant of EBOV from West Africa was distinct from previously isolated *ebolaviruses*. The data described in this project shows that infection with the Makona variant of EBOV causes similar effects to those elicited by other variants of EBOV, and the Makona variant does not display unique host response signatures.

7.6 Publications

Appendix III: Andrew Bosworth, Stuart D. Dowall, Natasha Y. Rickett et al: *A comparison of host gene expression signatures associated with infection in vitro by the Makona and Ecran (Mayinga) variants of Ebola virus.* Scientific Reports 01/2017

Appendix IV: Miles W Carroll, David A Matthews, Julian A Hiscox, … , Andrew Bosworth et al: *Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa.* Nature 06/2015;

7.7 Future Work

The proposed mechanism governing the relationship between NFκB pro-inflammatory activity and the replication of EBOV and RESTV requires further exploration. The next step is to selectively target host factors under the regulatory control of NFκB complexes to evaluate their importance in viral lifecycle, it is likely that critical host factors exist within this group of host proteins. The discovery that NFκB pro-inflammatory activity was induced differently by RESTV GP and EBOV GP begs the question of whether the observed differences in porcine cells were mechanistically related. The aim would be to treat porcine and human cell lines with GP from RESTV or EBOV and compare the TLR4 induction. It was hypothesised that RESTV would poorly induce TLR4 in human cells but efficiently induce reactivity in porcine cells.

The application of transcriptomics and proteomics to understand the Makona variant of EBOV is a demonstration of this approach to understand new and emerging viruses which present a threat to human health. If other viruses were discovered this use of technology would enable a rapid, informed, evidence-based risk assessment of the danger posed by emerging uncharacterised viruses and contributes to preparing for the WHO prioritised "Disease X".

References

1. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. Nature 473:337-342.
2. Report of an International C. 1978. Ebola haemorrhagic fever in Zaire, 1976. Bulletin of the World Health Organization 56:271-293.
3. Gordon Smith CE, Simpson DIH, Bowen ETW, Zlotnik I. FATAL HUMAN DISEASE FROM VERVET MONKEYS. The Lancet 290:1119-1121.
4. Luby JP, Sanders CV. 1969. Green monkey disease ("marburg virus" disease): A new zoonosis. Annals of Internal Medicine 71:657-660.
5. Kuhn JH, Bao Y, Bavari S, Becker S, Bradfute S, Brauburger K, Rodney Brister J, Bukreyev AA, Cai Y, Chandran K, Davey RA, Dolnik O, Dye JM, Enterlein S, Gonzalez JP, Formenty P, Freiberg AN, Hensley LE, Hoenen T, Honko AN, Ignatyev GM, Jahrling PB, Johnson KM, Klenk HD, Kobinger G, Lackemeyer MG, Leroy EM, Lever MS, Muhlberger E, Netesov SV, Olinger GG, Palacios G, Patterson JL, Paweska JT, Pitt L, Radoshitzky SR, Ryabchikova EI, Saphire EO, Shestopalov AM, Smither SJ, Sullivan NJ, Swanepoel R, Takada A, Towner JS, van der Groen G, Volchkov VE, Volchkova VA, Wahl-Jensen V, Warren TK, Warfield KL, et al. 2014. Virus nomenclature below the species level: a standardized nomenclature for filovirus strains and variants rescued from cDNA. Arch Virol 159:1229-37.
6. Beeching NJ, Fenech M, Houlihan CF. 2014. Ebola virus disease. BMJ : British Medical Journal 349.
7. Organisation WH. 2017. Ebola Fact Sheet. Accessed 17/07/17
8. Report of a WHOIST. 1978. Ebola haemorrhagic fever in Sudan, 1976. Bulletin of the World Health Organization 56:247-270.
9. Formenty P, Boesch C, Wyers M, Steiner C, Donati F, Dind F, Walker F, Le Guenno B. 1999. Ebola virus outbreak among wild chimpanzees living in a rain forest of Cote d'Ivoire. J Infect Dis 179 Suppl 1:S120-6.
10. Roddy P, Howard N, Van Kerkhove MD, Lutwama J, Wamala J, Yoti Z, Colebunders R, Palma PP, Sterk E, Jeffs B, Van Herp M, Borchert M. 2012. Clinical manifestations and case management of Ebola haemorrhagic fever caused by a newly identified virus strain, Bundibugyo, Uganda, 2007-2008. PLoS One 7:e52986.
11. Anonymous. 1989. Ebola virus infection in imported primates--Virginia, 1989. MMWR Morb Mortal Wkly Rep 38:831-2, 837-8.

12. Emond RT, Evans B, Bowen ET, Lloyd G. 1977. A case of Ebola virus infection. Br Med J 2:541-4.
13. Simpson DI. 1978. Viral haemorrhagic fevers of man. Bull World Health Organ 56:819-32.
14. Baskerville A, Fisher-Hoch SP, Neild GH, Dowsett AB. 1985. Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. J Pathol 147:199-209.
15. Anonymous. 1985. Viral haemorrhagic fevers. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser 721.
16. Muyembe-Tamfum JJ, Kipasa M, Kiyungu C, Colebunders R. 1999. Ebola Outbreak in Kikwit, Democratic Republic of the Congo: Discovery and Control Measures. The Journal of Infectious Diseases 179:S259-S262.
17. Hall RC, Hall RC, Chapman MJ. 2008. The 1995 Kikwit Ebola outbreak: lessons hospitals and physicians can apply to future viral epidemics. Gen Hosp Psychiatry 30:446-52.
18. Georges-Courbot MC, Lu CY, Lansoud-Soukate J, Leroy E, Baize S. 1997. Isolation and partial molecular characterisation of a strain of Ebola virus during a recent epidemic of viral haemorrhagic fever in Gabon. Lancet 349:181.
19. Kuhn JH, Lofts LL, Kugelman JR, Smither SJ, Lever MS, van der Groen G, Johnson KM, Radoshitzky SR, Bavari S, Jahrling PB, Towner JS, Nichol ST, Palacios G. 2014. Reidentification of Ebola Virus E718 and ME as Ebola Virus/H.sapiens-tc/COD/1976/Yambuku-Ecran. Genome Announc 2.
20. Carroll SA, Towner JS, Sealy TK, McMullan LK, Khristova ML, Burt FJ, Swanepoel R, Rollin PE, Nichol ST. 2013. Molecular evolution of viruses of the family Filoviridae based on 97 whole-genome sequences. J Virol 87:2608-16.
21. Kuhn JH, Dodd LE, Wahl-Jensen V, Radoshitzky SR, Bavari S, Jahrling PB. 2011. Evaluation of perceived threat differences posed by filovirus variants. Biosecur Bioterror 9:361-71.
22. Baize S, Pannetier D, Oestereich L, Rieger T, Koivogui L, Magassouba N, Soropogui B, Sow MS, Keita S, De Clerck H, Tiffany A, Dominguez G, Loua M, Traore A, Kolie M, Malano ER, Heleze E, Bocquin A, Mely S, Raoul H, Caro V, Cadar D, Gabriel M, Pahlmann M, Tappe D, Schmidt-Chanasit J, Impouma B, Diallo AK, Formenty P, Van Herp M, Gunther S. 2014. Emergence of Zaire Ebola virus disease in Guinea. N Engl J Med 371:1418-25.
23. Nanclares C, Kapetshi J, Lionetto F, de la Rosa O, Tamfun J-JM, Alia M, Kobinger G, Bernasconi A. 2016. Ebola Virus Disease, Democratic Republic of the Congo, 2014. Emerging Infectious Diseases 22:1579-1586.
24. Feldmann H, Geisbert TW. 2011. Ebola haemorrhagic fever. Lancet 377:849-62.
25. Cantoni D, Hamlet A, Michaelis M, Wass MN, Rossman JS. 2016. Risks Posed by Reston, the Forgotten Ebolavirus. mSphere 1.
26. Miranda ME, Miranda NL. 2011. Reston ebolavirus in humans and animals in the Philippines: a review. J Infect Dis 204 Suppl 3:S757-60.
27. Cornish JP, Diaz L, Ricklefs SM, Kanakabandi K, Sword J, Jahrling PB, Kuhn JH, Porcella SF, Johnson RF. 2017. Sequence of Reston Virus Isolate AZ-1435, an Ebolavirus Isolate Obtained during the 1989-1990 Reston Virus Epizootic in the United States. Genome Announc 5.
28. Taniguchi S, Sayama Y, Nagata N, Ikegami T, Miranda ME, Watanabe S, Iizuka I, Fukushi S, Mizutani T, Ishii Y, Saijo M, Akashi H, Yoshikawa Y, Kyuwa S,

Morikawa S. 2012. Analysis of the humoral immune responses among cynomolgus macaque naturally infected with Reston virus during the 1996 outbreak in the Philippines. BMC Vet Res 8:189.
29. Pan Y, Zhang W, Cui L, Hua X, Wang M, Zeng Q. 2014. Reston virus in domestic pigs in China. Archives of Virology 159:1129-1132.
30. CDC. 2017. Chronology of Ebola Virus Disease Outbreaks. Accessed 15/07/17
31. Swanepoel R, Leman PA, Burt FJ, Zachariades NA, Braack LE, Ksiazek TG, Rollin PE, Zaki SR, Peters CJ. 1996. Experimental inoculation of plants and animals with Ebola virus. Emerg Infect Dis 2:321-5.
32. Leroy E, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, Delicat A, Paweska J, Gonzalez J, Swanepoel R. 2005. Fruit bats as reservoirs of Ebola virus. Nature 438:575 - 576.
33. Snowden A. 1979. Marburg disease: the 20th Century. Zimbabwe Rhod Nurse:59-81.
34. Peterson AT, Samy AM. 2016. Geographic potential of disease caused by Ebola and Marburg viruses in Africa. Acta Tropica 162:114-124.
35. Bausch DG, Nichol ST, Muyembe-Tamfum JJ, Borchert M, Rollin PE, Sleurs H, Campbell P, Tshioko FK, Roth C, Colebunders R, Pirard P, Mardel S, Olinda LA, Zeller H, Tshomba A, Kulidri A, Libande ML, Mulangu S, Formenty P, Grein T, Leirs H, Braack L, Ksiazek T, Zaki S, Bowen MD, Smit SB, Leman PA, Burt FJ, Kemp A, Swanepoel R. 2006. Marburg Hemorrhagic Fever Associated with Multiple Genetic Lineages of Virus. New England Journal of Medicine 355:909-919.
36. Towner JS, Amman BR, Sealy TK, Carroll SAR, Comer JA, Kemp A, Swanepoel R, Paddock CD, Balinandi S, Khristova ML, Formenty PBH, Albarino CG, Miller DM, Reed ZD, Kayiwa JT, Mills JN, Cannon DL, Greer PW, Byaruhanga E, Farnon EC, Atimnedi P, Okware S, Katongole-Mbidde E, Downing R, Tappero JW, Zaki SR, Ksiazek TG, Nichol ST, Rollin PE. 2009. Isolation of Genetically Diverse Marburg Viruses from Egyptian Fruit Bats. PLoS Pathogens 5:e1000536.
37. Yuan J, Zhang Y, Li J, Zhang Y, Wang LF, Shi Z. 2012. Serological evidence of ebolavirus infection in bats, China. Virol J 9:236.
38. Jayme SI, Field HE, de Jong C, Olival KJ, Marsh G, Tagtag AM, Hughes T, Bucad AC, Barr J, Azul RR, Retes LM, Foord A, Yu M, Cruz MS, Santos IJ, Lim TMS, Benigno CC, Epstein JH, Wang L-F, Daszak P, Newman SH. 2015. Molecular evidence of Ebola Reston virus infection in Philippine bats. Virology Journal 12:107.
39. Breman JG, Johnson KM, van der Groen G, Robbins CB, Szczeniowski MV, Ruti K, Webb PA, Meier F, Heymann DL. 1999. A search for Ebola virus in animals in the Democratic Republic of the Congo and Cameroon: ecologic, virologic, and serologic surveys, 1979-1980. Ebola Virus Study Teams. J Infect Dis 179 Suppl 1:S139-47.
40. Pigott DM, Golding N, Mylne A, Huang Z, Henry AJ, Weiss DJ, Brady OJ, Kraemer MUG, Smith DL, Moyes CL, Bhatt S, Gething PW, Horby PW, Bogoch II, Brownstein JS, Mekaru SR, Tatem AJ, Khan K, Hay SI. 2014. Mapping the zoonotic niche of Ebola virus disease in Africa. eLife 3:e04395.
41. Olival JK, Hayman TD. 2014. Filoviruses in Bats: Current Knowledge and Future Directions. Viruses 6.
42. Luis A, Hayman D, O'Shea T, Cryan P, Gilbert A, Pulliam J, Mills J, Timonin M, Willis C, Cunningham A, Fooks A, Rupprecht C, Wood J, Webb C. 2013. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proc R Soc Lond B Biol Sci 280:20122753.

43. Wolfe ND, Daszak P, Kilpatrick AM, Burke DS. 2005. Bushmeat Hunting, Deforestation, and Prediction of Zoonotic Disease. Emerging Infectious Diseases 11:1822-1827.
44. Bausch DG, Schwarz L. 2014. Outbreak of Ebola Virus Disease in Guinea: Where Ecology Meets Economy. PLoS Neglected Tropical Diseases 8:e3056.
45. Bausch DG, Towner JS, Dowell SF, Kaducu F, Lukwiya M, Sanchez A, Nichol ST, Ksiazek TG, Rollin PE. 2007. Assessment of the risk of Ebola virus transmission from bodily fluids and fomites. J Infect Dis 196 Suppl 2:S142-7.
46. Zumbrun EE, Bloomfield HA, Dye JM, Hunter TC, Dabisch PA, Garza NL, Bramel NR, Baker RJ, Williams RD, Nichols DK, Nalca A. 2012. A Characterization of Aerosolized Sudan Virus Infection in African Green Monkeys, Cynomolgus Macaques, and Rhesus Macaques. Viruses 4:2115-2136.
47. Mekibib B, Ariën KK. 2016. Aerosol Transmission of Filoviruses. Viruses 8:148.
48. Rewar S, Mirdha D. 2014. Transmission of ebola virus disease: an overview. Ann Glob Health 80:444-51.
49. Tiffany A, Dalziel BD, Kagume Njenge H, Johnson G, Nugba Ballah R, James D, Wone A, Bedford J, McClelland A. 2017. Estimating the number of secondary Ebola cases resulting from an unsafe burial and risk factors for transmission during the West Africa Ebola epidemic. PLOS Neglected Tropical Diseases 11:e0005491.
50. Leroy EM, Rouquet P, Formenty P, Souquière S, Kilbourne A, Froment J-M, Bermejo M, Smit S, Karesh W, Swanepoel R, Zaki SR, Rollin PE. 2004. Multiple Ebola Virus Transmission Events and Rapid Decline of Central African Wildlife. Science 303:387.
51. Le Gouar PJ, Vallet D, David L, Bermejo M, Gatti S, Levréro F, Petit EJ, Ménard N. 2009. How Ebola Impacts Genetics of Western Lowland Gorilla Populations. PLOS ONE 4:e8375.
52. Bermejo M, Rodriguez-Teijeiro JD, Illera G, Barroso A, Vila C, Walsh PD. 2006. Ebola outbreak killed 5000 gorillas. Science 314:1564.
53. Formenty P, Hatz C, Le Guenno B, Stoll A, Rogenmoser P, Widmer A. 1999. Human infection due to Ebola virus, subtype Cote d'Ivoire: clinical and biologic presentation. J Infect Dis 179 Suppl 1:S48-53.
54. Reed PE, Mulangu S, Cameron KN, Ondzie AU, Joly D, Bermejo M, Rouquet P, Fabozzi G, Bailey M, Shen Z, Keele BF, Hahn B, Karesh WB, Sullivan NJ. 2014. A new approach for monitoring ebolavirus in wild great apes. PLoS Negl Trop Dis 8:e3143.
55. WHO. January 2015. One year into the Ebola outbreak: A deadly, tenacious and unforgiving virus. WHO Report.
56. Simon-Loriere E, Faye O, Faye O, Koivogui L, Magassouba N, Keita S, Thiberge JM, Diancourt L, Bouchier C, Vandenbogaert M, Caro V, Fall G, Buchmann JP, Matranga CB, Sabeti PC, Manuguerra JC, Holmes EC, Sall AA. 2015. Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic. Nature doi:10.1038/nature14612.
57. Team WHOER. 2014. Ebola virus disease in West Africa--the first 9 months of the epidemic and forward projections. N Engl J Med 371:1481-95.
58. Weidmann M, Sall AA, Manuguerra JC, Koivogui L, Adjami A, Traore FF, Hedlund KO, Lindegren G, Mirazimi A. 2011. Quantitative analysis of particles, genomes and infectious particles in supernatants of haemorrhagic fever virus cell cultures. Virol J 8:81.

59. Noda T, Sagara H, Suzuki E, Takada A, Kida H, Kawaoka Y. 2002. Ebola Virus VP40 Drives the Formation of Virus-Like Filamentous Particles Along with GP. Journal of Virology 76:4855-4865.
60. Martínez MJ, Biedenkopf N, Volchkova V, Hartlieb B, Alazard-Dany N, Reynard O, Becker S, Volchkov V. 2008. Role of Ebola Virus VP30 in Transcription Reinitiation. Journal of Virology 82:12569-12573.
61. Boehmann Y, Enterlein S, Randolf A, Muhlberger E. 2005. A reconstituted replication and transcription system for Ebola virus Reston and comparison with Ebola virus Zaire. Virology 332:406-17.
62. Muhlberger E, Weik M, Volchkov VE, Klenk HD, Becker S. 1999. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. J Virol 73:2333-42.
63. Zhang AP, Abelson DM, Bornholdt ZA, Liu T, Woods VL, Jr., Saphire EO. 2012. The ebolavirus VP24 interferon antagonist: know your enemy. Virulence 3:440-5.
64. Garcia-Dorival I, Wu W, Armstrong SD, Barr JN, Carroll MW, Hewson R, Hiscox JA. 2016. Elucidation of the Cellular Interactome of Ebola Virus Nucleoprotein and Identification of Therapeutic Targets. J Proteome Res 15:4290-4303.
65. Brauburger K, Boehmann Y, Tsuda Y, Hoenen T, Olejnik J, Schumann M, Ebihara H, Muhlberger E. 2014. Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. J Virol 88:12558-71.
66. Volchkova VA, Dolnik O, Martinez MJ, Reynard O, Volchkov VE. 2011. Genomic RNA Editing and Its Impact on Ebola Virus Adaptation During Serial Passages in Cell Culture and Infection of Guinea Pigs. Journal of Infectious Diseases 204:S941-S946.
67. Mehedi M, Falzarano D, Seebach J, Hu X, Carpenter MS, Schnittler HJ, Feldmann H. 2011. A new Ebola virus nonstructural glycoprotein expressed through RNA editing. J Virol 85:5406-14.
68. Mehedi M, Hoenen T, Robertson S, Ricklefs S, Dolan MA, Taylor T, Falzarano D, Ebihara H, Porcella SF, Feldmann H. 2013. Ebola virus RNA editing depends on the primary editing site sequence and an upstream secondary structure. PLoS Pathog 9:e1003677.
69. Volchkova VA, Dolnik O, Martinez MJ, Reynard O, Volchkov VE. 2011. Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs. J Infect Dis 204 Suppl 3:S941-6.
70. He J, Melnik LI, Komin A, Wiedman G, Fuselier T, Morris CF, Starr CG, Searson PC, Gallaher WR, Hristova K, Garry RF, Wimley WC. 2017. Ebola Virus Delta Peptide is a Viroporin. Journal of Virology doi:10.1128/jvi.00438-17.
71. Wahl-Jensen VM, Afanasieva TA, Seebach J, Stroher U, Feldmann H, Schnittler HJ. 2005. Effects of Ebola virus glycoproteins on endothelial cell activation and barrier function. J Virol 79:10442-50.
72. Hoenen T, Groseth A, de Kok-Mercado F, Kuhn JH, Wahl-Jensen V. 2011. Minigenomes, transcription and replication competent virus-like particles and beyond: reverse genetics systems for filoviruses and other negative stranded hemorrhagic fever viruses. Antiviral Res 91:195-208.
73. Mühlberger E. 2007. Filovirus replication and transcription. Future virology 2:205-215.
74. Hoenen T, Shabman RS, Groseth A, Herwig A, Weber M, Schudt G, Dolnik O, Basler CF, Becker S, Feldmann H. 2012. Inclusion bodies are a site of ebolavirus replication. J Virol 86:11779-88.

75. Watanabe S, Noda T, Kawaoka Y. 2006. Functional Mapping of the Nucleoprotein of Ebola Virus. Journal of Virology 80:3743-3751.
76. Noda T, Hagiwara K, Sagara H, Kawaoka Y. 2010. Characterization of the Ebola virus nucleoprotein–RNA complex. The Journal of General Virology 91:1478-1483.
77. Kirchdoerfer RN, Abelson DM, Li S, Wood MR, Saphire EO. 2015. Assembly of the Ebola virus nucleoprotein from a chaperoned VP35 complex. Cell reports 12:140-149.
78. Noda T, Kolesnikova L, Becker S, Kawaoka Y. 2011. The Importance of the NP: VP35 Ratio in Ebola Virus Nucleocapsid Formation. The Journal of Infectious Diseases 204:S878-S883.
79. Kirchdoerfer RN, Moyer CL, Abelson DM, Saphire EO. 2016. The Ebola Virus VP30-NP Interaction Is a Regulator of Viral RNA Synthesis. PLOS Pathogens 12:e1005937.
80. Biedenkopf N, Schlereth J, Grünweller A, Becker S, Hartmann RK. 2016. RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. Journal of Virology 90:7481-7496.
81. Hartman AL, Bird BH, Towner JS, Antoniadou ZA, Zaki SR, Nichol ST. 2008. Inhibition of IRF-3 activation by VP35 is critical for the high level of virulence of ebola virus. J Virol 82:2699-704.
82. Basler CF, Mikulasova A, Martinez-Sobrido L, Paragas J, Muhlberger E, Bray M, Klenk HD, Palese P, Garcia-Sastre A. 2003. The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. J Virol 77:7945-56.
83. Ilinykh PA, Lubaki NM, Widen SG, Renn LA, Theisen TC, Rabin RL, Wood TG, Bukreyev A. 2015. Different Temporal Effects of Ebola Virus VP35 and VP24 Proteins on Global Gene Expression in Human Dendritic Cells. Journal of Virology 89:7567-7583.
84. Moller-Tank S, Maury W. 2014. Phosphatidylserine receptors: enhancers of enveloped virus entry and infection. Virology 468-470:565-80.
85. Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, Kuehne AI, Kranzusch PJ, Griffin AM, Ruthel G, Dal Cin P, Dye JM, Whelan SP, Chandran K, Brummelkamp TR. 2011. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature 477:340-3.
86. Lee JE, Fusco ML, Hessell AJ, Oswald WB, Burton DR, Saphire EO. 2008. Structure of the Ebola virus glycoprotein bound to a human survivor antibody. Nature 454:177-182.
87. Takada A, Fujioka K, Tsuiji M, Morikawa A, Higashi N, Ebihara H, Kobasa D, Feldmann H, Irimura T, Kawaoka Y. 2004. Human Macrophage C-Type Lectin Specific for Galactose and N-Acetylgalactosamine Promotes Filovirus Entry. Journal of Virology 78:2943-2947.
88. Marzi A, Gramberg T, Simmons G, Möller P, Rennekamp AJ, Krumbiegel M, Geier M, Eisemann J, Turza N, Saunier B, Steinkasserer A, Becker S, Bates P, Hofmann H, Pöhlmann S. 2004. DC-SIGN and DC-SIGNR Interact with the Glycoprotein of Marburg Virus and the S Protein of Severe Acute Respiratory Syndrome Coronavirus. Journal of Virology 78:12090-12095.
89. Zhao D, Han X, Zheng X, Wang H, Yang Z, Liu D, Han K, Liu J, Wang X, Yang W, Dong Q, Yang S, Xia X, Tang L, He F. 2016. The Myeloid LSECtin Is a DAP12-Coupled Receptor That Is Crucial for Inflammatory Response Induced by Ebola Virus Glycoprotein. PLOS Pathogens 12:e1005487.

90. Schornberg KL, Shoemaker CJ, Dube D, Abshire MY, Delos SE, Bouton AH, White JM. 2009. α5β1-Integrin controls ebolavirus entry by regulating endosomal cathepsins. Proceedings of the National Academy of Sciences 106:8003-8008.
91. Hunt CL, Kolokoltsov AA, Davey RA, Maury W. 2011. The Tyro3 Receptor Kinase Axl Enhances Macropinocytosis of Zaire Ebolavirus. Journal of Virology 85:334-347.
92. Aleksandrowicz P, Marzi A, Biedenkopf N, Beimforde N, Becker S, Hoenen T, Feldmann H, Schnittler H-J. 2011. Ebola Virus Enters Host Cells by Macropinocytosis and Clathrin-Mediated Endocytosis. The Journal of Infectious Diseases 204:S957-S967.
93. Garver WS, Krishnan K, Gallagos JR, Michikawa M, Francis GA, Heidenreich RA. 2002. Niemann-Pick C1 protein regulates cholesterol transport to the trans-Golgi network and plasma membrane caveolae. Journal of Lipid Research 43:579-589.
94. de La Vega M-A, Wong G, Kobinger GP, Qiu X. 2015. The Multiple Roles of sGP in Ebola Pathogenesis. Viral Immunology 28:3-9.
95. Iwasa A, Shimojima M, Kawaoka Y. 2011. sGP serves as a structural protein in Ebola virus infection. J Infect Dis 204 Suppl 3:S897-903.
96. Sui J, Marasco WA. 2002. Evidence against Ebola Virus sGP Binding to Human Neutrophils by a Specific Receptor. Virology 303:9-14.
97. Hoenen T, Biedenkopf N, Zielecki F, Jung S, Groseth A, Feldmann H, Becker S. 2010. Oligomerization of Ebola virus VP40 is essential for particle morphogenesis and regulation of viral transcription. J Virol 84:7053-63.
98. Bavari S, Bosio CM, Wiegand E, Ruthel G, Will AB, Geisbert TW, Hevey M, Schmaljohn C, Schmaljohn A, Aman MJ. 2002. Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. J Exp Med 195:593-602.
99. Gomis-Ruth FX, Dessen A, Timmins J, Bracher A, Kolesnikowa L, Becker S, Klenk HD, Weissenhorn W. 2003. The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. Structure 11:423-33.
100. Watt A, Moukambi F, Banadyga L, Groseth A, Callison J, Herwig A, Ebihara H, Feldmann H, Hoenen T. 2014. A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. J Virol 88:10511-24.
101. Basler CF, Amarasinghe GK. 2009. Evasion of interferon responses by Ebola and Marburg viruses. J Interferon Cytokine Res 29:511-20.
102. Garcia-Dorival I, Wu W, Dowall S, Armstrong S, Touzelet O, Wastling J, Barr JN, Matthews D, Carroll M, Hewson R, Hiscox JA. 2014. Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host-cell protein function. J Proteome Res 13:5120-35.
103. Mateo M, Carbonnelle C, Reynard O, Kolesnikova L, Nemirov K, Page A, Volchkova VA, Volchkov VE. 2011. VP24 is a molecular determinant of Ebola virus virulence in guinea pigs. J Infect Dis 204 Suppl 3:S1011-20.
104. Dowall SD, Matthews DA, Garcia-Dorival I, Taylor I, Kenny J, Hertz-Fowler C, Hall N, Corbin-Lickfett K, Empig C, Schlunegger K, Barr JN, Carroll MW, Hewson R, Hiscox JA. 2014. Elucidating variations in the nucleotide sequence of Ebola virus associated with increasing pathogenicity. Genome Biol 15:540.
105. Pappalardo M, Julia M, Howard MJ, Rossman JS, Michaelis M, Wass MN. 2016. Conserved differences in protein sequence determine the human pathogenicity of Ebolaviruses. Sci Rep 6:23743.

106. Dilley KA, Voorhies AA, Luthra P, Puri V, Stockwell TB, Lorenzi H, Basler CF, Shabman RS. 2017. The Ebola virus VP35 protein binds viral immunostimulatory and host RNAs identified through deep sequencing. PLoS ONE 12:e0178717.
107. Messaoudi I, Amarasinghe GK, Basler CF. 2015. Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. Nature reviews Microbiology 13:663-676.
108. Knipe DM, Howley P. 2013. Fields Virology. Wolters Kluwer Health.
109. Bray M, Geisbert TW. 2005. Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. Int J Biochem Cell Biol 37:1560-6.
110. Geisbert TW, Hensley LE, Larsen T, Young HA, Reed DS, Geisbert JB, Scott DP, Kagan E, Jahrling PB, Davis KJ. 2003. Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. Am J Pathol 163:2347-70.
111. Dube D, Schornberg KL, Stantchev TS, Bonaparte MI, Delos SE, Bouton AH, Broder CC, White JM. 2008. Cell Adhesion Promotes Ebola Virus Envelope Glycoprotein-Mediated Binding and Infection. Journal of Virology 82:7238-7242.
112. Martinez O, Johnson JC, Honko A, Yen B, Shabman RS, Hensley LE, Olinger GG, Basler CF. 2013. Ebola virus exploits a monocyte differentiation program to promote its entry. J Virol 87:3801-14.
113. Kreuels B, Wichmann D, Emmerich P, Schmidt-Chanasit J, de Heer G, Kluge S, Sow A, Renne T, Gunther S, Lohse AW, Addo MM, Schmiedel S. 2014. A case of severe Ebola virus infection complicated by gram-negative septicemia. N Engl J Med 371:2394-401.
114. Hunt L, Gupta-Wright A, Simms V, Tamba F, Knott V, Tamba K, Heisenberg-Mansaray S, Tamba E, Sheriff A, Conteh S, Smith T, Tobin S, Brooks T, Houlihan C, Cummings R, Fletcher T. 2015. Clinical presentation, biochemical, and haematological parameters and their association with outcome in patients with Ebola virus disease: an observational cohort study. Lancet Infect Dis 15:1292-9.
115. Ilinykh PA, Lubaki NM, Widen SG, Renn LA, Theisen TC, Rabin RL, Wood TG, Bukreyev A. 2015. Different temporal effects of Ebola virus VP35 and VP24 proteins on the global gene expression in human dendritic cells. J Virol doi:10.1128/JVI.00924-15.
116. Mohamadzadeh M, Chen L, Schmaljohn AL. 2007. How Ebola and Marburg viruses battle the immune system. Nat Rev Immunol 7:556-567.
117. Wauquier N, Becquart P, Padilla C, Baize S, Leroy EM. 2010. Human fatal zaire ebola virus infection is associated with an aberrant innate immunity and with massive lymphocyte apoptosis. PLoS Negl Trop Dis 4.
118. Wahl-Jensen V, Kurz S, Feldmann F, Buehler LK, Kindrachuk J, DeFilippis V, da Silva Correia J, Fruh K, Kuhn JH, Burton DR, Feldmann H. 2011. Ebola virion attachment and entry into human macrophages profoundly effects early cellular gene expression. PLoS Negl Trop Dis 5:e1359.
119. Gupta M, Spiropoulou C, Rollin PE. 2007. Ebola virus infection of human PBMCs causes massive death of macrophages, CD4 and CD8 T cell sub-populations in vitro. Virology 364:45-54.
120. Edwards MR, Liu G, Mire CE, Sureshchandra S, Luthra P, Yen B, Shabman RS, Leung DW, Messaoudi I, Geisbert TW, Amarasinghe GK, Basler CF. 2016. Differential Regulation of Interferon Responses by Ebola and Marburg Virus VP35 Proteins. Cell Rep 14:1632-40.

121. Yen BC, Basler CF. 2016. Effects of Filovirus IFN Antagonists on Responses of Human Monocyte-Derived Dendritic Cells to RNA virus infection. J Virol doi:10.1128/jvi.00191-16.
122. Wahl-Jensen VM, Afanasieva TA, Seebach J, Ströher U, Feldmann H, Schnittler H-J. 2005. Effects of Ebola Virus Glycoproteins on Endothelial Cell Activation and Barrier Function. Journal of Virology 79:10442-10450.
123. Escudero-Pérez B, Volchkova VA, Dolnik O, Lawrence P, Volchkov VE. 2014. Shed GP of Ebola Virus Triggers Immune Activation and Increased Vascular Permeability. PLOS Pathogens 10:e1004509.
124. McElroy AK, Erickson BR, Flietstra TD, Rollin PE, Nichol ST, Towner JS, Spiropoulou CF. 2014. Ebola hemorrhagic Fever: novel biomarker correlates of clinical outcome. J Infect Dis 210:558-66.
125. Antoniak S, Mackman N. 2014. Multiple roles of the coagulation protease cascade during virus infection. Blood 123:2605-13.
126. Geisbert TW, Young HA, Jahrling PB, Davis KJ, Kagan E, Hensley LE. 2003. Mechanisms Underlying Coagulation Abnormalities in Ebola Hemorrhagic Fever: Overexpression of Tissue Factor in Primate Monocytes/Macrophages Is a Key Event. The Journal of Infectious Diseases 188:1618-1629.
127. Geisbert TW, Young HA, Jahrling PB, Davis KJ, Larsen T, Kagan E, Hensley LE. 2003. Pathogenesis of Ebola Hemorrhagic Fever in Primate Models : Evidence that Hemorrhage Is Not a Direct Effect of Virus-Induced Cytolysis of Endothelial Cells. The American Journal of Pathology 163:2371-2382.
128. Castell JV, Gómez-Lechón MJ, David M, Andus T, Geiger T, Trullenque R, Fabra R, Heinrich PC. 1989. Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. FEBS Letters 242:237-239.
129. Baize S, Leroy EM, Georges AJ, Georges-Courbot MC, Capron M, Bedjabaga I, Lansoud-Soukate J, Mavoungou E. 2002. Inflammatory responses in Ebola virus-infected patients. Clinical and Experimental Immunology 128:163-168.
130. Clark IA. 0000. The advent of the cytokine storm. Immunol Cell Biol 85:271-273.
131. Liu Q, Zhou Y-h, Yang Z-q. 2016. The cytokine storm of severe influenza and development of immunomodulatory therapy. Cellular and Molecular Immunology 13:3-10.
132. Ebihara H, Rockx B, Marzi A, Feldmann F, Haddock E, Brining D, LaCasse RA, Gardner D, Feldmann H. 2011. Host response dynamics following lethal infection of rhesus macaques with Zaire ebolavirus. J Infect Dis 204 Suppl 3:S991-9.
133. Tsalik EL, Langley RJ, Dinwiddie DL, Miller NA, Yoo B, van Velkinburgh JC, Smith LD, Thiffault I, Jaehne AK, Valente AM, Henao R, Yuan X, Glickman SW, Rice BJ, McClain MT, Carin L, Corey GR, Ginsburg GS, Cairns CB, Otero RM, Fowler VG, Jr., Rivers EP, Woods CW, Kingsmore SF. 2014. An integrated transcriptome and expressed variant analysis of sepsis survival and death. Genome Med 6:111.
134. Olejnik J, Alonso J, Schmidt KM, Yan Z, Wang W, Marzi A, Ebihara H, Yang J, Patterson JL, Ryabchikova E, Muhlberger E. 2013. Ebola virus does not block apoptotic signaling pathways. J Virol 87:5384-96.
135. Okumura A, Pitha PM, Yoshimura A, Harty RN. 2010. Interaction between Ebola virus glycoprotein and host toll-like receptor 4 leads to induction of proinflammatory cytokines and SOCS1. J Virol 84:27-33.
136. Kash JC, Muhlberger E, Carter V, Grosch M, Perwitasari O, Proll SC, Thomas MJ, Weber F, Klenk HD, Katze MG. 2006. Global suppression of the host antiviral

response by Ebola- and Marburgviruses: increased antagonism of the type I interferon response is associated with enhanced virulence. J Virol 80:3009-20.

137. Garamszegi S, Yen JY, Honko AN, Geisbert JB, Rubins KH, Geisbert TW, Xia Y, Hensley LE, Connor JH. 2014. Transcriptional correlates of disease outcome in anticoagulant-treated non-human primates infected with ebolavirus. PLoS Negl Trop Dis 8:e3061.

138. Canales RD, Luo Y, Willey JC, Austermiller B, Barbacioru CC, Boysen C, Hunkapiller K, Jensen RV, Knight CR, Lee KY, Ma Y, Maqsodi B, Papallo A, Peters EH, Poulter K, Ruppel PL, Samaha RR, Shi L, Yang W, Zhang L, Goodsaid FM. 2006. Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotechnol 24.

139. Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY, Luo Y, Sun YA, Willey JC, Setterquist RA, Fischer GM, Tong W, Dragan YP, Dix DJ, Frueh FW, Goodsaid FM, Herman D, Jensen RV, Johnson CD, Lobenhofer EK, Puri RK, Schrf U, Thierry-Mieg J, Wang C, Wilson M, Wolber PK, Zhang L, Amur S, Bao W, Barbacioru CC, Lucas AB, Bertholet V, Boysen C, Bromley B, Brown D, Brunner A, Canales R, Cao XM, Cebula TA, Chen JJ, Cheng J, Chu TM, Chudin E, Corson J, Corton JC, Croner LJ, et al. 2006. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24.

140. Robinson M, McCarthy D, Smyth G. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England) 26:139-140.

141. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley D, Pimentel H, Salzberg S, Rinn J, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protocols 7:562-578.

142. Connor JH, Yen J, Caballero IS, Garamszegi S, Malhotra S, Lin K, Hensley L, Goff AJ. 2015. Transcriptional Profiling of the Immune Response to Marburg Virus Infection. Journal of Virology doi:10.1128/jvi.01142-15.

143. Cilloniz C, Ebihara H, Ni C, Neumann G, Korth MJ, Kelly SM, Kawaoka Y, Feldmann H, Katze MG. 2011. Functional genomics reveals the induction of inflammatory response and metalloproteinase gene expression during lethal Ebola virus infection. J Virol 85:9060-8.

144. Timmins J, Schoehn G, Ricard-Blum S, Scianimanico S, Vernet T, Ruigrok RW, Weissenhorn W. 2003. Ebola virus matrix protein VP40 interaction with human cellular factors Tsg101 and Nedd4. J Mol Biol 326:493-502.

145. Yasuda J, Nakao M, Kawaoka Y, Shida H. 2003. Nedd4 regulates egress of Ebola virus-like particles from host cells. J Virol 77:9987-92.

146. Lu J, Qu Y, Liu Y, Jambusaria R, Han Z, Ruthel G, Freedman BD, Harty RN. 2013. Host IQGAP1 and Ebola Virus VP40 Interactions Facilitate Virus-Like Particle Egress. Journal of Virology 87:7777-7780.

147. Takahashi K, Halfmann P, Oyama M, Kozuka-Hata H, Noda T, Kawaoka Y. 2013. DNA Topoisomerase 1 Facilitates the Transcription and Replication of the Ebola Virus Genome. Journal of Virology 87:8862-8869.

148. Yuan S, Cao L, Ling H, Dang M, Sun Y, Zhang X, Chen Y, Zhang L, Su D, Wang X, Rao Z. 2015. TIM-1 acts a dual-attachment receptor for Ebolavirus by interacting directly with viral GP and the PS on the viral envelope. Protein & Cell 6:814-824.

149. Luthra P, Jordan DS, Leung DW, Amarasinghe GK, Basler CF. 2015. Ebola Virus VP35 Interaction with Dynein LC8 Regulates Viral RNA Synthesis. Journal of Virology 89:5148-5153.
150. Shabman RS, Leung DW, Johnson J, Glennon N, Gulcicek EE, Stone KL, Leung L, Hensley L, Amarasinghe GK, Basler CF. 2011. DRBP76 Associates With Ebola Virus VP35 and Suppresses Viral Polymerase Function. The Journal of Infectious Diseases 204:S911-S918.
151. Wynne J, Shiell B, Marsh G, Boyd V, Harper J, Heesom K, Monaghan P, Zhou P, Payne J, Klein R, Todd S, Mok L, Green D, Bingham J, Tachedjian M, Baker M, Matthews D, Wang L-F. 2014. Proteomics informed by transcriptomics reveals Hendra virus sensitizes bat cells to TRAIL-mediated apoptosis. Genome Biology 15:532.
152. Venkataramanan KP, Min L, Hou S, Jones SW, Ralston MT, Lee KH, Papoutsakis ET. 2015. Complex and extensive post-transcriptional regulation revealed by integrative proteomic and transcriptomic analysis of metabolite stress response in Clostridium acetobutylicum. Biotechnology for Biofuels 8:81.
153. Lackner DH, Schmidt MW, Wu S, Wolf DA, Bähler J. 2012. Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. Genome Biology 13:R25-R25.
154. Reid SP, Leung LW, Hartman AL, Martinez O, Shaw ML, Carbonnelle C, Volchkov VE, Nichol ST, Basler CF. 2006. Ebola Virus VP24 Binds Karyopherin α1 and Blocks STAT1 Nuclear Accumulation. Journal of Virology 80:5156-5167.
155. Reid SP, Shurtleff AC, Costantino JA, Tritsch SR, Retterer C, Spurgers KB, Bavari S. 2014. HSPA5 is an essential host factor for Ebola virus infection. Antiviral Res 109:171-4.
156. Spurgers KB, Alefantis T, Peyser BD, Ruthel GT, Bergeron AA, Costantino JA, Enterlein S, Kota KP, Boltz RC, Aman MJ, Delvecchio VG, Bavari S. 2010. Identification of essential filovirion-associated host factors by serial proteomic analysis and RNAi screen. Mol Cell Proteomics 9:2690-703.
157. Randall RE, Goodbourn S. 2008. Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. Journal of General Virology 89:1-47.
158. Spiropoulou CF, Ranjan P, Pearce MB, Sealy TK, Albariño CG, Gangappa S, Fujita T, Rollin PE, Nichol ST, Ksiazek TG, Sambhara S. 2009. RIG-I activation inhibits ebolavirus replication. Virology 392:11-15.
159. Deguine J, Barton GM. 2014. MyD88: a central player in innate immune signaling. F1000Prime Reports 6:97.
160. Leung DW, Shabman RS, Farahbakhsh M, Prins KC, Borek DM, Wang T, Muhlberger E, Basler CF, Amarasinghe GK. 2010. Structural and functional characterization of Reston Ebola virus VP35 interferon inhibitory domain. J Mol Biol 399:347-57.
161. Harcourt BH, Sanchez A, Offermann MK. 1998. Ebola virus inhibits induction of genes by double-stranded RNA in endothelial cells. Virology 252:179-88.
162. Gantke T, Boussouf S, Janzen J, Morrice NA, Howell S, Muhlberger E, Ley SC. 2013. Ebola virus VP35 induces high-level production of recombinant TPL-2-ABIN-2-NF-kappaB1 p105 complex in co-transfected HEK-293 cells. Biochem J 452:359-65.
163. Martinez O, Valmas C, Basler CF. 2007. Ebola virus-like particle-induced activation of NF-kappaB and Erk signaling in human dendritic cells requires the glycoprotein mucin domain. Virology 364:342-54.

164. Gibb TR, Norwood DA, Jr., Woollen N, Henchal EA. 2002. Viral replication and host gene expression in alveolar macrophages infected with Ebola virus (Zaire strain). Clin Diagn Lab Immunol 9:19-27.
165. Lawrence T. 2009. The Nuclear Factor NF-κB Pathway in Inflammation. Cold Spring Harbor Perspectives in Biology 1:a001651.
166. Baetu T, Kwon H, Sharma S, Grandvaux N, Hiscott J. 2001. Disruption of NF-kappa B signaling reveals a novel role for NF-kappa B in the regulation of TNF-related apoptosis-inducing ligand expression. J Immunol 167:3164 - 3173.
167. Wek RC, Jiang HY, Anthony TG. 2006. Coping with stress: eIF2 kinases and translational control. Biochem Soc Trans 34:7-11.
168. Shrestha N, Bahnan W, Wiley DJ, Barber G, Fields KA, Schesser K. 2012. Eukaryotic Initiation Factor 2 (eIF2) Signaling Regulates Proinflammatory Cytokine Expression and Bacterial Invasion. Journal of Biological Chemistry 287:28738-28744.
169. Jindal S, Malkovsky M. Stress responses to viral infection. Trends in Microbiology 2:89-91.
170. Protter DSW, Parker R. Principles and Properties of Stress Granules. Trends in Cell Biology 26:668-679.
171. Serfling E, Berberich-Siebelt F, Avots A, Chuvpilo S, Klein-Hessling S, Jha MK, Kondo E, Pagel P, Schulze-Luehrmann J, Palmetshofer A. 2004. NFAT and NF-kappaB factors-the distant relatives. Int J Biochem Cell Biol 36:1166-70.
172. Christian F, Smith EL, Carmody RJ. 2016. The Regulation of NF-κB Subunits by Phosphorylation. Cells 5:12.
173. Sun S-C. 2012. The noncanonical NF-κB pathway. Immunological Reviews 246:125-140.
174. Hiscott J, Kwon H, Génin P. 2001. Hostile takeovers: viral appropriation of the NF-kB pathway. Journal of Clinical Investigation 107:143-151.
175. Medzhitov R. 2001. Toll-like receptors and innate immunity. Nat Rev Immunol 1:135-45.
176. Geisbert TW, Strong JE, Feldmann H. 2015. Considerations in the Use of Nonhuman Primate Models of Ebola Virus and Marburg Virus Infection. J Infect Dis 212 Suppl 2:S91-7.
177. Hoenen T, Marzi A, Scott DP, Feldmann F, Callison J, Safronetz D, Ebihara H, Feldmann H. 2015. Soluble Glycoprotein Is Not Required for Ebola Virus Virulence in Guinea Pigs. J Infect Dis 212 Suppl 2:S242-6.
178. Genin M, Clement F, Fattaccioli A, Raes M, Michiels C. 2015. M1 and M2 macrophages derived from THP-1 cells differentially modulate the response of cancer cells to etoposide. BMC Cancer 15:577.
179. Qin Z. 2012. The use of THP-1 cells as a model for mimicking the function and regulation of monocytes and macrophages in the vasculature. Atherosclerosis 221:2-11.
180. Daigneault M, Preston JA, Marriott HM, Whyte MK, Dockrell DH. 2010. The identification of markers of macrophage differentiation in PMA-stimulated THP-1 cells and monocyte-derived macrophages. PLoS One 5:e8668.
181. Reed, Muench H. 1938. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS. Am J Epidemiol 27:493-497.
182. Trombley AR, Wachter L, Garrison J, Buckley-Beason VA, Jahrling J, Hensley LE, Schoepp RJ, Norwood DA, Goba A, Fair JN, Kulesh DA. 2010. Comprehensive Panel of Real-Time TaqMan™ Polymerase Chain Reaction Assays for Detection and

Absolute Quantification of Filoviruses, Arenaviruses, and New World Hantaviruses. The American Journal of Tropical Medicine and Hygiene 82:954-960.
183. Girod M, Biarc J, Enjalbert Q, Salvador A, Antoine R, Dugourd P, Lemoine J. 2014. Implementing visible 473 nm photodissociation in a Q-Exactive mass spectrometer: towards specific detection of cysteine-containing peptides. Analyst 139:5523-5530.
184. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:10.
185. Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9:357-359.
186. Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotech 26:1367-1372.
187. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. Journal of Proteome Research 10:1794-1805.
188. Kerber R, Krumkamp R, Diallo B, Jaeger A, Rudolf M, Lanini S, Bore JA, Koundouno FR, Becker-Ziaja B, Fleischmann E, Stoecker K, Meschi S, Mely S, Newman EN, Carletti F, Portmann J, Korva M, Wolff S, Molkenthin P, Kis Z, Kelterbaum A, Bocquin A, Strecker T, Fizet A, Castilletti C, Schudt G, Ottowell L, Kurth A, Atkinson B, Badusche M, Cannas A, Pallasch E, Bosworth A, Yue C, Palyi B, Ellerbrok H, Kohl C, Oestereich L, Logue CH, Ludtke A, Richter M, Ngabo D, Borremans B, Becker D, Gryseels S, Abdellati S, Vermoesen T, Kuisma E, Kraus A, Liedigk B, et al. 2016. Analysis of Diagnostic Findings From the European Mobile Laboratory in Gueckedou, Guinea, March 2014 Through March 2015. J Infect Dis doi:10.1093/infdis/jiw269.
189. Hofmann-Winkler H, Gnirß K, Wrensch F, Pöhlmann S. 2015. Comparative Analysis of Host Cell Entry of Ebola Virus From Sierra Leone, 2014, and Zaire, 1976. The Journal of Infectious Diseases 212:S172-S180.
190. Urbanowicz Richard A, McClure C P, Sakuntabhai A, Sall Amadou A, Kobinger G, Müller Marcel A, Holmes Edward C, Rey Félix A, Simon-Loriere E, Ball Jonathan K. 2016. Human Adaptation of Ebola Virus during the West African Outbreak. Cell 167:1079-1087.e5.
191. Andrea M, Friederike F, Patrick WH, Dana PS, Stephan G, Heinz F. 2015. Delayed Disease Progression in Cynomolgus Macaques Infected with Ebola Virus Makona Strain. Emerging Infectious Disease journal 21.
192. Dove BK, Surtees R, Bean TJH, Munday D, Wise HM, Digard P, Carroll MW, Ajuh P, Barr JN, Hiscox JA. 2012. A quantitative proteomic analysis of lung epithelial (A549) cells infected with 2009 pandemic influenza A virus using stable isotope labelling with amino acids in cell culture. PROTEOMICS 12:1431-1436.
193. Carroll MW, Matthews DA, Hiscox JA, Elmore MJ, Pollakis G, Rambaut A, Hewson R, Garcia-Dorival I, Bore JA, Koundouno R, Abdellati S, Afrough B, Aiyepada J, Akhilomen P, Asogun D, Atkinson B, Badusche M, Bah A, Bate S, Baumann J, Becker D, Becker-Ziaja B, Bocquin A, Borremans B, Bosworth A, Boettcher JP, Cannas A, Carletti F, Castilletti C, Clark S, Colavita F, Diederich S, Donatus A, Duraffour S, Ehichioya D, Ellerbrok H, Fenandez-Garcia MD, Fizet A, Fleischmann E, Gryseels S, Hermelink A, Hinzmann J, Hopf-Guevara U, Ighodalo Y, Jameson L, Kelterbaum A, Kis Z, Kloth S, Kohl C, Korva M, et al. 2015. Temporal and spatial

analysis of the 2014-2015 Ebola virus outbreak in West Africa. Nature doi:10.1038/nature14594.

194. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouedraogo N, Afrough B, Bah A, Baum JH, Becker-Ziaja B, Boettcher JP, Cabeza-Cabrerizo M, Camino-Sanchez A, Carter LL, Doerrbecker J, Enkirch T, Dorival IG, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou LE, Koropogui M, Kosgey A, Kuisma E, Logue CH, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallasch E, Patrono LV, Portmann J, Repits JG, Rickett NY, Sachse A, Singethan K, Vitoriano I, Yemanaberhan RL, Zekeng EG, Racine T, Bello A, Sall AA, Faye O, et al. 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature doi:10.1038/nature16996.

195. Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, Kanneh L, Jalloh S, Momoh M, Fullah M, Dudas G, Wohl S, Moses LM, Yozwiak NL, Winnicki S, Matranga CB, Malboeuf CM, Qu J, Gladden AD, Schaffner SF, Yang X, Jiang P-P, Nekoui M, Colubri A, Coomber MR, Fonnie M, Moigboi A, Gbakie M, Kamara FK, Tucker V, Konuwa E, Saffa S, Sellu J, Jalloh AA, Kovoma A, Koninga J, Mustapha I, Kargbo K, Foday M, Yillah M, Kanneh F, Robert W, Massally JLB, Chapman SB, Bochicchio J, Murphy C, Nusbaum C, Young S, Birren BW, Grant DS, Scheiffelin JS, et al. 2014. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science 345:1369-1372.

196. Towner JS, Rollin PE, Bausch DG, Sanchez A, Crary SM, Vincent M, Lee WF, Spiropoulou CF, Ksiazek TG, Lukwiya M, Kaducu F, Downing R, Nichol ST. 2004. Rapid Diagnosis of Ebola Hemorrhagic Fever by Reverse Transcription-PCR in an Outbreak Setting and Assessment of Patient Viral Load as a Predictor of Outcome. Journal of Virology 78:4330-4341.

197. Bell A, Lewandowski K, Myers R, Wooldridge D, Aarons E, Simpson A, Vipond R, Jacobs M, Gharbia S, Zambon M. 2015. Genome sequence analysis of Ebola virus in clinical samples from three British healthcare workers, August 2014 to March 2015. Euro Surveill 20.

198. Read JM, Diggle PJ, Chirombo J, Solomon T, Baylis M. 2015. Effectiveness of screening for Ebola at airports. The Lancet 385:23-24.

199. Jacobs M, Rodger A, Bell DJ, Bhagani S, Cropley I, Filipe A, Gifford RJ, Hopkins S, Hughes J, Jabeen F, Johannessen I, Karageorgopoulos D, Lackenby A, Lester R, Liu RS, MacConnachie A, Mahungu T, Martin D, Marshall N, Mepham S, Orton R, Palmarini M, Patel M, Perry C, Peters SE, Porter D, Ritchie D, Ritchie ND, Seaton RA, Sreenu VB, Templeton K, Warren S, Wilkie GS, Zambon M, Gopal R, Thomson EC. 2016. Late Ebola virus relapse causing meningoencephalitis: a case report. Lancet 388:498-503.

200. Velasquez GE, Aibana O, Ling EJ, Diakite I, Mooring EQ, Murray MB. 2015. Time From Infection to Disease and Infectiousness for Ebola Virus Disease, a Systematic Review. Clin Infect Dis doi:10.1093/cid/civ531.

201. Kuhn JH. 2008. Filoviruses. A compendium of 40 years of epidemiological, clinical, and laboratory studies. Arch Virol Suppl 20:13-360.

202. Xuan Liu ES, Cesar Munoz-Fontela, Sam Haldenby, Natasha Y. Rickett, Isabel Garcia-Dorival, Yongxiang Fang, Yper Hall, Elsa-Gayle Zekeng, Anja Ludtke, Dong Xia, Romy Kerber, Ralf Krumkamp, Sophie Duraffour, Daouda Sissoko, John Kenny, Nichola Rockliffe, E. Diane Williamson, Thomas R. Laws, Magassouba N'Faly, David

A. Matthews, Stephan Gunther, Andrew R. Cossins, Armand Sprecher, John H. Connor, Miles W. Carroll and Julian A. Hiscox. 2016. Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus. Genome Biology Pre-Print.

203. Kash JC, Mühlberger E, Carter V, Grosch M, Perwitasari O, Proll SC, Thomas MJ, Weber F, Klenk H-D, Katze MG. 2006. Global Suppression of the Host Antiviral Response by Ebola- and Marburgviruses: Increased Antagonism of the Type I Interferon Response Is Associated with Enhanced Virulence. Journal of Virology 80:3009-3020.

204. Barrenas F, Green RR, Thomas MJ, Law GL, Proll SC, Engelmann F, Messaoudi I, Marzi A, Feldmann H, Katze MG. 2015. Next-Generation Sequencing Reveals a Controlled Immune Response to Zaire Ebola Virus Challenge in Cynomolgus Macaques Immunized with Vesicular Stomatitis Virus Expressing Zaire Ebola Virus Glycoprotein (VSVΔG/EBOVgp). Clinical and Vaccine Immunology 22:354-356.

205. Rasmussen AL, Okumura A, Ferris MT, Green R, Feldmann F, Kelly SM, Scott DP, Safronetz D, Haddock E, LaCasse R, Thomas MJ, Sova P, Carter VS, Weiss JM, Miller DR, Shaw GD, Korth MJ, Heise MT, Baric RS, de Villena FP-M, Feldmann H, Katze MG. 2014. Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. Science 346:987-991.

206. Noyori O, Nakayama E, Maruyama J, Yoshida R, Takada A. 2013. Suppression of Fas-mediated apoptosis via steric shielding by filovirus glycoproteins. Biochem Biophys Res Commun 441:994-8.

207. Sanchez A, Ksiazek TG, Rollin PE, Miranda ME, Trappier SG, Khan AS, Peters CJ, Nichol ST. 1999. Detection and molecular characterization of Ebola viruses causing disease in human and nonhuman primates. J Infect Dis 179 Suppl 1:S164-9.

208. Guito JC, Albarino CG, Chakrabarti AK, Towner JS. 2017. Novel activities by ebolavirus and marburgvirus interferon antagonists revealed using a standardized in vitro reporter system. Virology 501:147-165.

209. Gibb TR, Bray M, Geisbert TW, Steele KE, Kell WM, Davis KJ, Jaax NK. 2001. Pathogenesis of experimental Ebola Zaire virus infection in BALB/c mice. J Comp Pathol 125:233-42.

210. Traore K, Trush MA, George M, Spannhake EW, Anderson W, Asseffa A. 2005. Signal transduction of phorbol 12-myristate 13-acetate (PMA)-induced growth inhibition of human monocytic leukemia THP-1 cells is reactive oxygen dependent. Leuk Res 29.

211. Maess MB, Sendelbach S, Lorkowski S. 2010. Selection of reliable reference genes during THP-1 monocyte differentiation into macrophages. BMC Mol Biol 11.

212. Nagahara Y, Nagamori T, Tamegai H, Hitokuwada M, Yoshimi Y, Ikekita M, Shinomiya T. 2011. Inulin stimulates phagocytosis of PMA-treated THP-1 macrophages by involvement of PI3-kinases and MAP kinases. Biofactors 37.

213. Chen R-F, Wang L, Cheng J-T, Yang KD. 2012. Induction of IFNα or IL-12 depends on differentiation of THP-1 cells in dengue infections without and with antibody enhancement. BMC Infectious Diseases 12:340.

214. Chareonsirisuthigul T, Kalayanarooj S, Ubol S. 2007. Dengue virus (DENV) antibody-dependent enhancement of infection upregulates the production of anti-inflammatory cytokines, but suppresses anti-DENV free radical and pro-inflammatory cytokine production, in THP-1 cells. J Gen Virol 88.

215. Bremner TA, D'Costa N, Dickson LA, Asseffa A. 1996. A decrease in glucose 6-phosphate dehydrogenase activity and mRNA is an early event in phorbol ester-induced differentiation of thp-1 promonocytic leukemia cells. Life Sci 58.
216. Barilli A, Rotoli BM, Visigalli R, Bussolati O, Gazzola GC, Dall'Asta V. 2011. Arginine transport in human monocytic leukemia THP-1 cells during macrophage differentiation. J Leukoc Biol 90.
217. Nelder JA, Wedderburn RWM. 1972. Generalized Linear Models. Journal of the Royal Statistical Society Series A (General) 135:370-384.
218. Liu X, Speranza E, Munoz-Fontela C, Haldenby S, Rickett NY, Garcia-Dorival I, Fang Y, Hall Y, Zekeng EG, Ludtke A, Xia D, Kerber R, Krumkamp R, Duraffour S, Sissoko D, Kenny J, Rockliffe N, Williamson ED, Laws TR, N'Faly M, Matthews DA, Gunther S, Cossins AR, Sprecher A, Connor JH, Carroll MW, Hiscox JA. 2017. Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus. Genome Biol 18:4.
219. Kwon AT, Arenillas DJ, Worsley Hunt R, Wasserman WW. 2012. oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. G3 (Bethesda) 2:987-1002.
220. Olejnik J, Forero A, Deflubé LR, Hume AJ, Manhart WA, Nishida A, Marzi A, Katze MG, Ebihara H, Rasmussen AL, Mühlberger E. 2017. Ebolaviruses associated with differential pathogenicity induce distinct host responses in human macrophages. Journal of Virology doi:10.1128/jvi.00179-17.
221. Pierre ER. 2009. Filoviruses: A Compendium of 40 Years of Epidemiological, Clinical, and Laboratory Studies. Emerging Infectious Disease journal 15:2079.
222. Rhein BA, Powers LS, Rogers K, Anantpadma M, Singh BK, Sakurai Y, Bair T, Miller-Hunt C, Sinn P, Davey RA, Monick MM, Maury W. 2015. Interferon-? Inhibits Ebola Virus Infection. PLoS Pathog 11:e1005263.
223. Strong J, Wong G, Jones S, Grolla A, Theriault S, Kobinger G, Feldmann H. 2008. Stimulation of Ebola virus production from persistent infection through activation of the Ras/MAPK pathway. Proc Natl Acad Sci U S A 105:17982 - 17987.
224. Wrensch F, Karsten CB, Gnirß K, Hoffmann M, Lu K, Takada A, Winkler M, Simmons G, Pöhlmann S. 2015. Interferon-Induced Transmembrane Protein–Mediated Inhibition of Host Cell Entry of Ebolaviruses. Journal of Infectious Diseases doi:10.1093/infdis/jiv255.
225. Lee H, Nishiura H. 2017. Recrudescence of Ebola virus disease outbreak in West Africa, 2014-2016. Int J Infect Dis 64:90-92.
226. Andrea M, Friederike F, Patrick WH, Dana PS, Stephan G, Heinz F. 2015. Delayed Disease Progression in Cynomolgus Macaques Infected with Ebola Virus Makona Strain. Emerging Infectious Disease journal 21:1777.
227. Grolla A, Jones SM, Fernando L, Strong JE, Stroher U, Moller P, Paweska JT, Burt F, Pablo Palma P, Sprecher A, Formenty P, Roth C, Feldmann H. 2011. The use of a mobile laboratory unit in support of patient management and epidemiological surveillance during the 2005 Marburg Outbreak in Angola. PLoS Negl Trop Dis 5:e1183.
228. Maganga GD, Kapetshi J, Berthet N, Kebela Ilunga B, Kabange F, Mbala Kingebeni P, Mondonge V, Muyembe JJ, Bertherat E, Briand S, Cabore J, Epelboin A, Formenty P, Kobinger G, Gonzalez-Angulo L, Labouba I, Manuguerra JC, Okwo-Bele JM, Dye C, Leroy EM. 2014. Ebola virus disease in the Democratic Republic of Congo. N Engl J Med 371:2083-91.

229. Nanyonga M, Saidu J, Ramsay A, Shindo N, Bausch DG. 2015. Sequelae of Ebola Virus Disease, Kenema District, Sierra Leone. Clinical Infectious Diseases doi:10.1093/cid/civ795.
230. Bosworth A, Dowall SD, Garcia-Dorival I, Rickett NY, Bruce CB, Matthews DA, Fang Y, Aljabr W, Kenny J, Nelson C, Laws TR, Williamson ED, Stewart JP, Carroll MW, Hewson R, Hiscox JA. 2017. A comparison of host gene expression signatures associated with infection in vitro by the Makona and Ecran (Mayinga) variants of Ebola virus. Sci Rep 7:43144.
231. Ludwig S, Planz O. 2008. Influenza viruses and the NF-kappaB signaling pathway - towards a novel concept of antiviral therapy. Biol Chem 389:1307-12.
232. Pinto R, Herold S, Cakarova L, Hoegner K, Lohmeyer J, Planz O, Pleschka S. 2011. Inhibition of influenza virus-induced NF-kappaB and Raf/MEK/ERK activation can reduce both virus titers and cytokine expression simultaneously in vitro and in vivo. Antiviral Res 92:45-56.
233. Kumar N, Xin Z-t, Liang Y, Ly H, Liang Y. 2008. NF-κB Signaling Differentially Regulates Influenza Virus RNA Synthesis. Journal of Virology 82:9880-9889.
234. Dove BK, Surtees R, Bean TJ, Munday D, Wise HM, Digard P, Carroll MW, Ajuh P, Barr JN, Hiscox JA. 2012. A quantitative proteomic analysis of lung epithelial (A549) cells infected with 2009 pandemic influenza A virus using stable isotope labelling with amino acids in cell culture. Proteomics 12:1431-6.
235. Munday DC, Emmott E, Surtees R, Lardeau CH, Wu W, Duprex WP, Dove BK, Barr JN, Hiscox JA. 2010. Quantitative proteomic analysis of A549 cells infected with human respiratory syncytial virus. Mol Cell Proteomics 9:2438-59.
236. Bian T, Gibbs JD, Orvell C, Imani F. 2012. Respiratory syncytial virus matrix protein induces lung epithelial cell cycle arrest through a p53 dependent pathway. PLoS One 7:e38052.
237. Dunham EC, Banadyga L, Groseth A, Chiramel AI, Best SM, Ebihara H, Feldmann H, Hoenen T. 2015. Assessing the contribution of interferon antagonism to the virulence of West African Ebola viruses. Nature Communications 6:8000.
238. Aljabr W, Touzelet O, Pollakis G, Wu W, Munday DC, Hughes M, Hertz-Fowler C, Kenny J, Fearns R, Barr JN, Matthews DA, Hiscox JA. 2016. Investigating the Influence of Ribavirin on Human Respiratory Syncytial Virus RNA Synthesis by Using a High-Resolution Transcriptome Sequencing Approach. J Virol 90:4876-88.
239. Vanden Berghe W, Sabbe L, Kaileh M, Haegeman G, Heyninck K, 2012. Molecular insight in the multifunctional activities of Withaferin A. Biochem Pharm 84:1282-1291

Appendix I: Consent and Permissions

The following pages display the consent to use patient samples for a case study presented in Chapter 3 of this thesis. Shown are approval letters, participant information forms and signed consent forms. Also shown are permission letters from collaborators for the use of images used in the Thesis introduction

NHS
Health Research Authority
National Research Ethics Service
NRES Committee Northwest – Greater Manchester West
3rd Floor
Barlow House
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7434

07 September 2015

Mr Andrew Bosworth
NIHR HPRU PhD Student
Public Health England
Porton Down
SP4 0JG

Dear Mr Bosworth

Study title: **Host Transcriptomic Changes During Clinical Infection With Ebola Virus**
REC reference: **15/NW/0744**
IRAS project ID: **189909**

Thank you for your submission. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 07 September 2015

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Participant consent form	2.0	05 September 2015
Participant information sheet (PIS)	2.0	05 September 2015

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	*Version*	*Date*
IRAS Checklist XML [Checklist_03092015]		03 September 2015
Letter from sponsor [Sponsorship In Principle Letter]	1	02 September 2015
Participant consent form	2.0	05 September 2015
Participant information sheet (PIS)	2.0	05 September 2015
REC Application Form [REC_Form_03092015]		03 September 2015
Research protocol or project proposal [Study Overview]	1	27 August 2015
Summary CV for Chief Investigator (CI) [Chief Investigator CV]	1	02 September 2015
Summary CV for supervisor (student research) [Supervisor CV]	1	02 September 2015

You should ensure that the sponsor has a copy of the final documentation for the study. It is the

Public Health England

Participant Information Sheet

Dear Pauline Cafferky

You are being invited to be part of a study which aims to improve our understanding of Ebola Virus Disease, to help us determine the changes which take place in a patient during the course of infection, and see whether they are related to the severity of symptoms. This work will enable the discovery of biomarkers and may help in the development of treatments for viral haemorrhagic fevers such as Ebola Virus. This study will also aim to improve our understanding of the lasting effects of Ebola Virus in survivors of the disease, with the aim of improving patient management for the large numbers of survivors in West Africa.

Why have I been chosen for this study?

During your recent Ebola illness and after recovery you had samples taken as part of your clinical care. These sequential samples provide a unique opportunity to analyse the changes that occur in patients and how they respond to the virus during the course of the disease.

What will participation involve?

We are asking you to donate the samples of blood and joint aspirate which were taken for your clinical care for research purposes. No additional samples other than those which have been taken for your clinical care, or which will be taken as part of your clinical follow-up, will be required. However, we need your consent before these can be used for research. We also ask that anonymised information from your medical records is made available to the researchers to help them interpret the research findings. Since your samples provide a unique set of clinical samples from an Ebola patient in the UK, we very much hope that you will agree to this.

Do I have to take part?

You do not have to donate your samples for research and if you choose not to do so, your clinical care will not be affected.

What are the risks of taking part?

There are no additional risks since we are only asking to use samples that have already been taken or which will be taken in the future as part of the clinical management and follow-up of your Ebola illness.

What are the possible benefits of taking part?

There will be no direct benefits to yourself. This study will not reveal anything of clinical significance to you but it may help future patients.

Will my taking part in this study be kept confidential?

Your samples will be anonymised and study staff will not have access to your personal details. Relevant information from your medical records will be anonymised and provided to the study staff in response to specific questions about your illness. Any publications arising from this research will not include your personal details, but because there have been very few cases of Ebola in the UK and there has been much media coverage of cases, it is possible that someone might deduce your identity. Only analysed data relevant to the study will be published or disseminated.

Who is organising and funding the research?

This study is sponsored by Public Health England, and funded by the National Institute for Health Research, Health Protection Research Unit for Emerging and Zoonotic Infections. It has undergone ethical review by the Greater Manchester West Research Ethics Committee.

Does this research involve genetic testing?

Next-generation sequencing data will be analysed for gene expression analysis related to Ebola infection and will not be used to identify other infectious organisms other than Ebola Virus, nor will this be used for any form of genetic analysis.

Who can I contact for further information?

We very much hope that you will agree to donate your samples for this study. If you would like more information about the study, or have any questions that have not been answered, please contact your local clinical care team in the first instance, who will forward questions to the Public Health England, Porton Down for response.

Public Health England

Host Transcriptomic Changes During Clinical Infection With Ebola Virus

Reference: 189909 05/09/2015 Version 2.0

CONSENT FORM

Title of Project: **Host Transcriptomic Changes During Clinical Infection With Ebola Virus**
Chief Investigator: Mr. Andrew Bosworth MRes BSc

Please initial all boxes

1. I confirm that I have read and understand the information sheet (reference 189909 05/09/2015 version 2.0) for the above project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. [AL]

2. I understand that my participation in this study is voluntary. [AL]

3. I agree to take part in the above project, giving consent for samples collected and sent to Public Health England, Porton Down as part of diagnosis, monitoring and follow-up related or thought to be related to my prior infection with Ebola Virus, to be used for the purposes described in the study information sheet (reference 189909 05/09/2015 version 2.0). [AL]

4. I understand that I can withdraw from this project at any time until my sample is anonymised and delivered to research study scientists. [AL]

5. I agree that samples may be stored in the receiving clinical pathology laboratory until October 2017 for further research purposes subject to ethical permission being granted. [AL]

6. I consent to clinical case notes and data relevant to my prior infection with Ebola virus and continuing symptoms suspected to be related to this previous infection, being seen by the study research team. [AL]

7. I consent to the results of the research being published as a case report in relevant journals and scientific publications. [AL]

PAULINE CAMPBELL Name of Participant	8/9/15 Date	 Signature
TERESA TOWN Person taking consent	8/9/15 Date	 Signature

Message from Dr. Thomas Strecker of the University of Marburg an der Lahn, giving permission for the use of cell culture images produced by Dr. Verena Krahling.

From: Thomas Strecker
Sent: 30 August 2017 13:33
To: Andrew Bosworth
Cc: Verena Krähling
Subject: Re: Question for Strecker.

Hi Andrew,
As the pictures were provided by Verena, it is Okay to acknowledge only her.
Best regards and good luck with your thesis
Thomas

Sent from mobile device

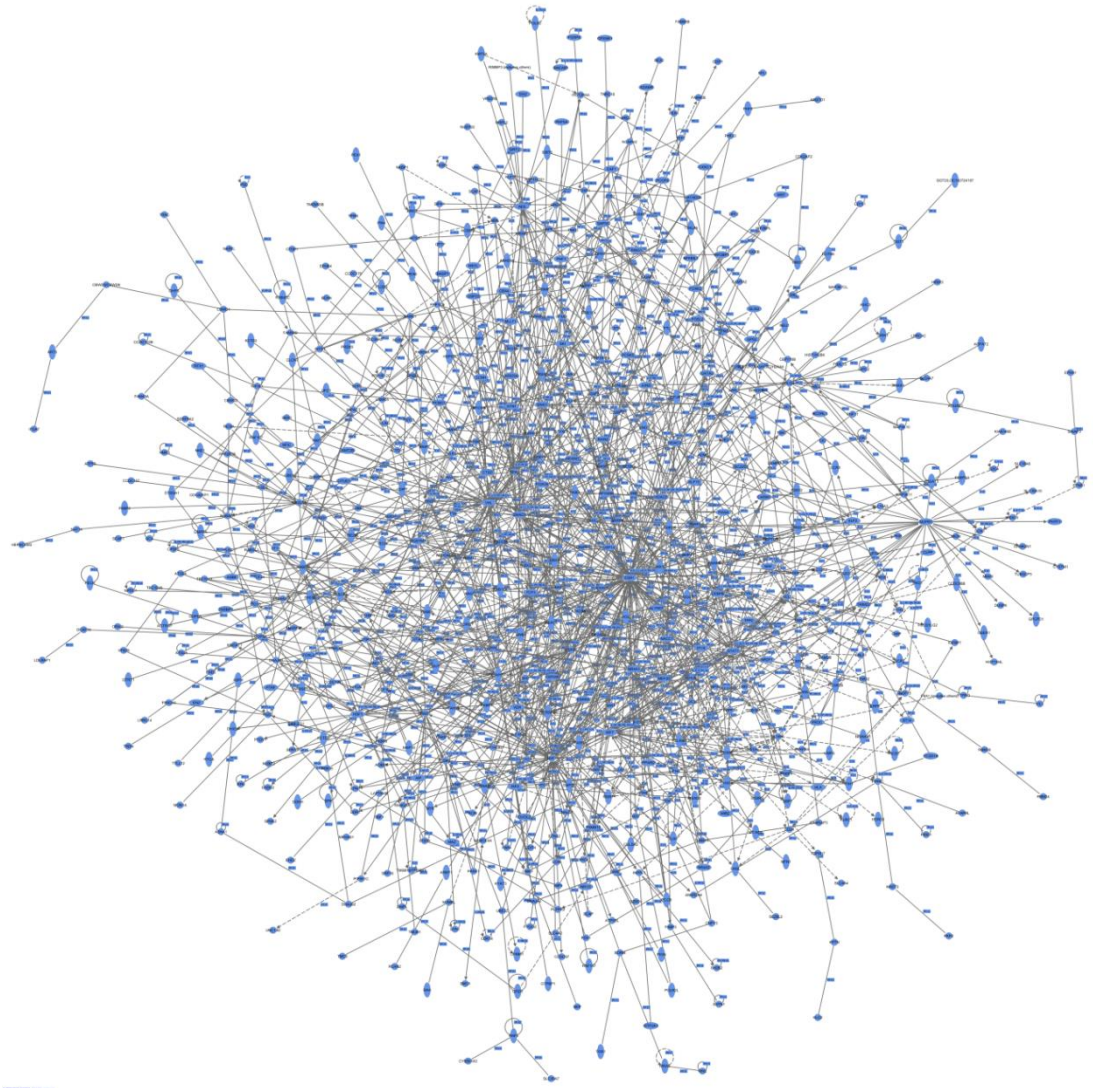
Message from Dr. Matthew Hannah giving permission for the use of TEM images of EBOV, Makona.

From: Matthew Hannah
Sent: 31 May 2017 07:36
To: Andrew Bosworth
Subject: RE: EM images of Ebola

No problem Andrew, the images are very much a team effort. The virus was grown up in HCM by Gopal et al. I took the images on the EM. The images were then colourized by Science Photolibrary (SPL). All the colour Ebola images are joint copyright between PHE and SPL. However, we (i.e. PHE) are allowed to use them for our internal publications (I'm pretty sure that a PhD thesis will count as such)

Dr Matthew Hannah PhD

Appendix II: Supplementary Figure



Supplementary Figure: Complete expanded map illustrating all indirect and direct interacting transcripts identified in a transcriptomic study of THP-1 derived macrophages infected with EBOV, Makona variant. 653 transcripts are shown; image produced using Qiagen's Ingenuity Pathway Analysis software.

Appendix III: Associated Publication

www.nature.com/scientificreports

OPEN

Received: 08 July 2016
Accepted: 18 January 2017
Published: 27 February 2017

A comparison of host gene expression signatures associated with infection *in vitro* by the Makona and Ecran (Mayinga) variants of Ebola virus

Andrew Bosworth[1,2], Stuart D. Dowall[1,2], Isabel Garcia-Dorival[3], Natasha Y. Rickett[2,3], Christine B. Bruce[4], David A. Matthews[5], Yongxiang Fang[6], Waleed Aljabr[3], John Kenny[6], Charlotte Nelson[6], Thomas R. Laws[7], E. Diane Williamson[7], James P. Stewart[3], Miles W. Carroll[1,2], Roger Hewson[1,2] & Julian A. Hiscox[2,3]

The Ebola virus (EBOV) variant Makona (which emerged in 2013) was the causative agent of the largest outbreak of Ebola Virus Disease recorded. Differences in virus-host interactions between viral variants have potential consequences for transmission, disease severity and mortality. A detailed profile of the cellular changes induced by the Makona variant compared with other Ebola virus variants was lacking. In this study, A549 cells, a human cell line with a robust innate response, were infected with the Makona variant or with the Ecran variant originating from the 1976 outbreak in Central Africa. The abundance of viral and cellular mRNA transcripts was profiled using RNASeq and differential gene expression analysis performed. Differences in effects of each virus on the expression of interferon-stimulated genes were also investigated in A549 NPro cells where the type 1 interferon response had been attenuated. Cellular transcriptomic changes were compared with those induced by human respiratory syncytial virus (HRSV), a virus with a similar genome organisation and replication strategy to EBOV. Pathway and gene ontology analysis revealed differential expression of functionally important genes; including genes involved in the inflammatory response, cell proliferation, leukocyte extravasation and cholesterol biosynthesis. Whilst there was overlap with HRSV, there was unique commonality to the EBOV variants.

The evolution of Ebolaviruses of varying pathogenicity complicates assessment of the public health significance of novel filoviruses. Ebolaviruses are a diverse genus within the Filoviridae family. Five species have been characterised; from those non-pathogenic in humans to those with high case-fatality rates (CFR) ranging from 25–90%[1,2]. The recently isolated Ebola Virus (EBOV) Makona variant, from a clinical case of Ebola Virus Disease (EVD) in Macenta, Guinea, has been observed to have a varying CFR of 70% to 50% at later stages of the outbreak[1]. Detailed analysis revealed non-synonymous changes in the genome of the Makona variant compared with other previously identified variants of EBOV[3,4], though the significance of many of these phenotypic changes is poorly understood.

Work *in vivo* using *Macaca fascicularis* as an animal model identified a potential shift in the pathogenesis of infection with EBOV Makona compared to other variants. Results showed a reduced pace of disease progression correlating with viral titre[5]. Differences exist in clinical data concerning disease course and mortality rates

[1]National Infection Service, Public Health England, Porton Down, Salisbury, UK. [2]National Institute of Health Research, Health Protection Research Unit in Emerging and Zoonotic Infections, Liverpool, UK. [3]Institute of Infection and Global Health, University of Liverpool, UK. [4]High Containment Microbiology, Public Health England, Porton Down, Salisbury, UK. [5]School of Molecular and Cellular Medicine, University of Bristol, UK. [6]Centre for Genomics Research, University of Liverpool, UK. [7]Defence Science and Technology Laboratory, Porton Down, UK. Correspondence and requests for materials should be addressed to R.H. (email: roger.hewson@phe.gov.uk) or J.A.H. (email: julian.hiscox@liverpool.ac.uk)

Appendix IV: Associated Publication

LETTER

OPEN

doi:10.1038/nature14594

Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa

Miles W. Carroll[1,2,3], David A. Matthews[4]*, Julian A. Hiscox[5]*, Michael J. Elmore[1]*, Georgios Pollakis[5]*, Andrew Rambaut[6,7,8]*, Roger Hewson[1,2,9], Isabel García-Dorival[5], Joseph Akoi Bore[2,10,11], Raymond Koundouno[2,10,11], Said Abdellati[2,12], Babak Afrough[1,2], John Aiyepada[2,13], Patience Akhilomen[2,13], Danny Asogun[2,13], Barry Atkinson[1,2], Marlis Badusche[2,14,15], Amadou Bah[2,16], Simon Bate[1,2], Jan Baumann[2,14], Dirk Becker[2,15,17], Beate Becker-Ziaja[2,14,15], Anne Bocquin[2,18,19], Benny Borremans[2,20], Andrew Bosworth[1,2,5], Jan Peter Boettcher[2,21], Angela Cannas[2,22], Fabrizio Carletti[2,22], Concetta Castilletti[2,22], Simon Clark[1,2], Francesca Colavita[2,22], Sandra Diederich[2,15,23], Adomeh Donatus[2,13], Sophie Duraffour[2,14,24], Deborah Ehichioya[2,14,25], Heinz Ellerbrok[2,21], Maria Dolores Fernandez-Garcia[2,26], Alexandra Fizet[2,18,27], Erna Fleischmann[2,15,28], Sophie Gryseels[2,20], Antje Hermelink[2,21], Julia Hinzmann[2,21], Ute Hopf-Guevara[2,21], Yemisi Ighodalo[2,13], Lisa Jameson[1,2], Anne Kelterbaum[2,15,17], Zoltan Kis[2,29], Stefan Kloth[2,21], Claudia Kohl[2,21], Miša Korva[2,30], Annette Kraus[2,31], Eeva Kuisma[1,2], Andreas Kurth[2,21], Britta Liedigk[2,14,15], Christopher H. Logue[1,2], Anja Lüdtke[2,15,32], Piet Maes[2,24], James McCowen[1,2], Stéphane Mély[2,18,19], Marc Mertens[2,15,23], Silvia Meschi[2,22], Benjamin Meyer[2,15,33], Janine Michel[2,21], Peter Molkenthin[2,15,28], César Muñoz-Fontela[2,15,32], Doreen Muth[2,15,33], Edmund N. C. Newman[1,2], Didier Ngabo[1,2], Lisa Oestereich[2,14,15], Jennifer Okosun[2,13], Thomas Olokor[2,13], Racheal Omiunu[2,13], Emmanuel Omomoh[2,13], Elisa Pallasch[2,14,15], Bernadett Pályi[2,29], Jasmine Portmann[2,34], Thomas Pottage[1,2], Catherine Pratt[1,2], Simone Priesnitz[2,35], Serena Quartu[2,22], Julie Rappe[2,36], Johanna Repits[2,37], Martin Richter[2,21], Martin Rudolf[2,14,15], Andreas Sachse[2,21], Kristina Maria Schmidt[2,21], Gordian Schudt[2,15,17], Thomas Strecker[2,15,17], Ruth Thom[1,2], Stephen Thomas[1,2], Ekaete Tobin[2,13], Howard Tolley[1,2], Jochen Trautner[2,38], Tine Vermoesen[2,12], Inês Vitoriano[1,2], Matthias Wagner[2,15,28], Svenja Wolff[2,15,17], Constanze Yue[2,21], Maria Rosaria Capobianchi[2,22], Birte Kretschmer[39], Yper Hall[1], John G. Kenny[40], Natasha Y. Rickett[5], Gytis Dudas[6], Cordelia E. M. Coltart[41], Romy Kerber[2,14,15], Damien Steer[42], Callum Wright[43], Francis Senyah[1], Sakoba Keita[44], Patrick Drury[45], Boubacar Diallo[46], Hilde de Clerck[47], Michel Van Herp[47], Armand Sprecher[47], Alexis Traore[48], Mandiou Diakite[49], Mandy Kader Konde[50], Lamine Koivogui[11], N'Faly Magassouba[10], Tatjana Avšič-Županc[2,30], Andreas Nitsche[2,21], Marc Strasser[2,34], Giuseppe Ippolito[2,22], Stephan Becker[2,15,17], Kilian Stoecker[2,15,28], Martin Gabriel[2,14,15], Hervé Raoul[2,19], Antonino Di Caro[2,22], Roman Wölfel[2,15,28], Pierre Formenty[45] & Stephan Günther[2,14,15]*

West Africa is currently witnessing the most extensive Ebola virus (EBOV) outbreak so far recorded[1–3]. Until now, there have been 27,013 reported cases and 11,134 deaths. The origin of the virus is thought to have been a zoonotic transmission from a bat to a two-year-old boy in December 2013 (ref. 2). From this index case the virus was spread by human-to-human contact throughout Guinea, Sierra Leone and Liberia. However, the origin of the particular virus in each country and time of transmission is not known and currently relies on epidemiological analysis, which may be unreliable owing to the difficulties of obtaining patient information. Here we trace the genetic evolution of EBOV in the current outbreak that has resulted in multiple lineages. Deep sequencing of 179 patient samples processed by the European Mobile Laboratory, the first diagnostics unit to be deployed to the epicentre of the outbreak in Guinea, reveals an epidemiological and evolutionary history of the epidemic from March 2014 to January 2015. Analysis of EBOV genome evolution has also benefited from a similar sequencing effort of patient samples from Sierra Leone. Our results confirm that the EBOV from Guinea moved into Sierra Leone, most likely in April or early May. The viruses of the Guinea/Sierra Leone lineage mixed around June/July 2014. Viral sequences covering August, September and October 2014 indicate that this lineage evolved independently within Guinea. These data can be used in conjunction with epidemiological information to test retrospectively the effectiveness of control measures, and provides an unprecedented window into the evolution of an ongoing viral haemorrhagic fever outbreak.

We used a deep sequencing approach to gain insight into the evolution of Ebola virus (EBOV) in Guinea from the ongoing West African outbreak. This was an approach based on analysis pipelines developed

[1]Public Health England, Porton Down, Wiltshire SP4 0JG, UK. [2]The European Mobile Laboratory Consortium, Bernhard-Nocht-Institute for Tropical Medicine, D-20359 Hamburg, Germany. [3]University of Southampton, South General Hospital, Southampton SO16 6YD, UK. [4]Department of Cellular and Molecular Medicine, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, UK. [5]Institute of Infection and Global Health, University of Liverpool, Liverpool L69 2BE, UK. [6]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 2FL, UK. [7]Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892, USA. [8]Centre for Immunology, Infection and Evolution, University of Edinburgh, Edinburgh EH9 2FL, UK. [9]London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. [10]Université Gamal Abdel Nasser de Conakry, Laboratoire des Fièvres Hémorragiques en Guinée, Conakry, Guinea. [11]Institut National de Santé Publique, Conakry, Guinea. [12]Institute of Tropical Medicine, B-2000 Antwerp, Belgium. [13]Institute of Lassa Fever Research and Control, Irrua Specialist Teaching Hospital, Irrua, Edo State, Nigeria. [14]Bernhard Nocht Institute for Tropical Medicine, D-20359 Hamburg, Germany. [15]German Centre for Infection Research (DZIF), 38124 Braunschweig, Germany. [16]Swiss Tropical and Public Health Institute, University of Basel, CH-4002 Basel, Switzerland. [17]Institute of Virology, Philipps University Marburg, 35043 Marburg, Germany. [18]National Reference Center for Viral Hemorrhagic Fevers, 69365 Lyon, France. [19]Laboratoire P4 Inserm-Jean Mérieux, US003 Inserm, 69365 Lyon, France. [20]Department of Biology, University of Antwerp, B-2020 Antwerp, Belgium. [21]Robert Koch Institute, 13353 Berlin, Germany. [22]National Institute for Infectious Diseases (INMI) Lazzaro Spallanzani, 00149 Rome, Italy. [23]Friedrich Loeffler Institute, Federal Research Institute for Animal Health, 17493 Greifswald, Insel Riems, Germany. [24]KU Leuven Rega institute, B-3000 Leuven, Belgium. [25]Redeemer's University, Osun State, Nigeria. [26]Centro Nacional de Microbiologia, Instituto de Salud Carlos III, 28029 Madrid, Spain. [27]Unité de Biologie des Infections Virales Emergentes, Institut Pasteur, 69365 Lyon, France. [28]Bundeswehr Institute of Microbiology, 80937 Munich, Germany. [29]National Center for Epidemiology, National Biosafety Laboratory, H-1097 Budapest, Hungary. [30]Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, SI-1000 Ljubljana, Slovenia. [31]Public Health Agency of Sweden, 171 82 Solna, Sweden. [32]Heinrich Pette Institute – Leibniz Institute for Experimental Virology, 20251 Hamburg, Germany. [33]Institute of Virology, University of Bonn, 53127 Bonn, Germany. [34]Federal Office for Civil Protection, Spiez Laboratory, CH-3700 Spiez, Switzerland. [35]Bundeswehr Hospital, 22049 Hamburg, Germany. [36]Institute of Virology and Immunology, CH-3147 Mittelhäusern, Switzerland. [37]Janssen-Cilag, SE-192 07 Sollentuna, Sweden. [38]Thünen Institute, D-22767 Hamburg, Germany. [39]Eurice - European Research and Project Office GmbH, 10115 Berlin, Germany. [40]Centre for Genomic Research, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK. [41]Department of Infection and Population Health, University College London, London WC1E 6JB, UK. [42]Research IT, University of Bristol, Bristol BS8 1HH, UK. [43]Advanced Computing Research Centre, University of Bristol, Bristol BS8 1HH, UK. [44]Ministry of Health Guinea, Conakry, Guinea. [45]World Health Organization, 1211 Geneva 27, Switzerland. [46]World Health Organization, Conakry, Guinea. [47]Médecins Sans Frontières, B-1050 Brussels, Belgium. [48]Section Prévention et Lutte contre la Maladie à la Direction Préfectorale de la Santé de Guéckédou, Guéckédou, Guinea. [49]Université Gamal Abdel Nasser de Conakry, CHU Donka, Conakry, Guinea. [50]Health and Sustainable Development Foundation, Conakry, Guinea.
*These authors contributed equally to this work.